Newcastle PRES 2011: Results and analysis

1 Introduction

In this document we outline the main findings of the 2011 Postgraduate Research Experience Survey (PRES) at Newcastle, which went "live" on March 1st and closed on May 17th. Of the 2,137 students invited to complete the survey, 926 responded, giving an overall response rate of just over 43%. Excluding 7 students who failed to provide their School/Institute, for each of the three Faculties there were 312/658, 322/765 and 285/714 students completing the PRES, giving individual Faculty response rates of about 47%, 42% and 40% for the Faculty of Medical Sciences (FMS), Humanities and Social Sciences (HaSS) and Science, Agriculture and Engineering (SAgE), respectively.

1.1 Aims of this report

Excluding questions asking for demographic information about the student, there were 13 main sections to the survey (see section 1.2 below). This report will:

1. give Faculty– and School/Institute–level comparisons of student responses to questions in each of these areas;
2. compare this year's results at Newcastle with results from previous years, examining the year–on–year performance of Schools/Institutes/Faculties;
3. draw comparisons between Newcastle's results in 2011 and those of other groups (e.g. Newcastle versus Russell Group);
4. consider an adjustment to this year's results which reduces bias as a result of differential response rates for different groups of students within each School/Institute.

1.2 Areas covered by the PRES

Of the 13 main sections to the PRES, 12 sections elicited student opinions numerically (the other section was for 'free response' comments from the responder relating to their overall experience as a postgraduate researcher at Newcastle; we do not consider these here). Throughout this report, we label these as:

1. Supervision
2. Skills Development
3. Infrastructure
4. Intellectual Climate
5. Goals & Standards
6. Professional Development
7. Roles/Responsibilities
8. Teaching Opportunities
9. Thesis examination
10. Expectations of various factors
11. Importance of various factors
12. Personal factors

For all but questions in section 10, students were asked to respond on a Likert scale: they were asked to rate, on a scale of 1 (strongly disagree) to 5 (strongly agree), their level of agreement

with several statements (where, mostly, 1 gives the most negative outcome and 5 gives the most favourable outcome) associated with each area. A different scale was used to gauge "expectations" in section 10, and the direction of the Likert scale varied for questions in section 12 (here, for some questions a score of 5 corresponds to the most positive outcome whilst for others it corresponds to the most negative).

1.3 Further information

Key findings will be highlighted throughout the report. Any detailed mathematical/statistical development will be clearly labelled as such, and can be omitted without any loss of clarity in the discussion. Before reading the main content of this report, the reader should also familiarise themselves with the following information.

1.3.1 The 'key set' and question 15g

We provide detailed analyses of responses to questions in sections **1**–**8** of the PRES; throughout this document, we refer to these as the *key set*. We consider these sections together as responses here were required on exactly the same Likert scale (see section 1.2 above).

Given that only 15 of the questionnaire respondents had actually taken their *viva voce* examination at the time of the PRES, we consider results from section **9** only very briefly. sections **10** and **11** ask students about their expectations and perceived importance of various factors, and so we consider responses to questions in these two sections together. In particular, it has been suggested that one question in section 10 be used to rate a responder's overall experience as a research student, in terms of how this has met with their expectations:

"*Please rate the overall experience of your research programme in terms of how it has met with your expectations (–3 = it is much more negative, 0 = it has met my expectations, +3 = it is much more positive).*"

In the questionnaire this was *question 15g*; thus, in most sections of this report we also give special attention to responses to this question. Responses to questions in section **12** are considered separately due to the slightly different nature of the Likert scale used (see above).

1.3.2 Aggregating scores

To summarise student responses at the School/Institute or Faculty level for each section in the key set, we average scores for each student across questions within each of these sections. For example, for each student i in School/Institute j we have a response string $\mathbf{R}_{i,j}$, e.g. $\mathbf{R}_{i,j} = (1, 2, 5, 4, \ldots)$, for each of the eight sections in the key set. For example, students are asked to rate their level of agreement to six statements about "Supervision" (section 1); for student 2 in School/Institute 13, this might give $\mathbf{R}_{2,13} = (4, 4, 5, 4, 3, 5)$. To summarise a student's response to each section, we then find the mean response for that student, $\bar{R}_{i,j}$; in our example this gives $\bar{R}_{2,13} = 4.167$. We can then summarise student responses for "Supervision" for this School/Institute as a whole by finding the mean or perhaps the median of $\bar{R}_{i,j}$ for all students in that School/Institute.

2 Newcastle PRES 2011: Internal comparisons

The aim of this section is to compare Schools/Institutes with: (i) each other; (ii) their Faculties, and (iii) the University as a whole. We also compare Faculties with: (i) each other, and (ii) the University as a whole. Simple graphical and numerical summaries by School/Institute, Faculty and the University as a whole will be presented and explained, as well as the results from more formal statistical analyses; some key messages, as suggested by the data, will be highlighted where appropriate.

2.1 Simple graphical/numerical summaries

2.1.1 The key set

We initially consider student responses to questions in the key set of the PRES. Students' responses are aggregated as explained in section 1.3.2, giving an overall summary – $\bar{R}_{i,j}$ – for each student i in School/Institute j for every section in the key set. We can then summarise student responses for each section in the key set at a School/Institute level by finding the mean or perhaps the median of $\bar{R}_{i,j}$ across all students in each School/Institute; dispersion can be measured using the standard deviation or by inspection of the quartiles of $\bar{R}_{i,j}$. We can, of course, collect Schools/Institutes within Faculties and summarise $\bar{R}_{i,j}$ across each of FMS, HaSS and SAgE; University metrics can also be obtained by summarising across all respondents together. A full set of such numerical results is given in Tables 5–7 in the Appendix.

Figure 1 shows box–and–whisker plots of $\bar{R}_{i,j}$ for each School/Institute j. Each plot shows the median and quartiles (the 'box'); 'whiskers' extend the full range of the sample (excluding outliers). Numbers on the horizontal axes correspond to the following (with the number of respondents given in parentheses):

FMS (312)	**HaSS** (322)	**SAgE** (285)
1. IAH (32)	**10.** ECLS (70)	**19.** AFRD (22)
2. ICMB (39)	**11.** NUBS (38)	**20.** BIOL (20)
3. ICM (50)	**12.** SACS (31)	**21.** CEAM (46)
4. IHG (26)	**13.** SAPL (31)	**22.** CHEM (14)
5. IHS (32)	**14.** SELS (79)	**23.** CIVG (48)
6. ION (23)	**15.** SGPS (34)	**24.** COMP (29)
7. MSGS (74)	**16.** SHIS (19)	**25.** EECE (56)
8. NICR (26)	**17.** SLAW (10)	**26.** MAST (19)
9. PSYC (10)	**18.** SMLS (10)	**27.** MATH (12)
		28. MECH (13)
		29. NIRES (6)

Some key points from Figure 1 and the tables of numerical results in the Appendix are given below/overleaf:

Supervision: Very little difference between Faculty averages; many School/Institute averages above Faculty/University averages; results from SLAW seem convincing, with even the lower quartile exceeding Faculty/University averages.

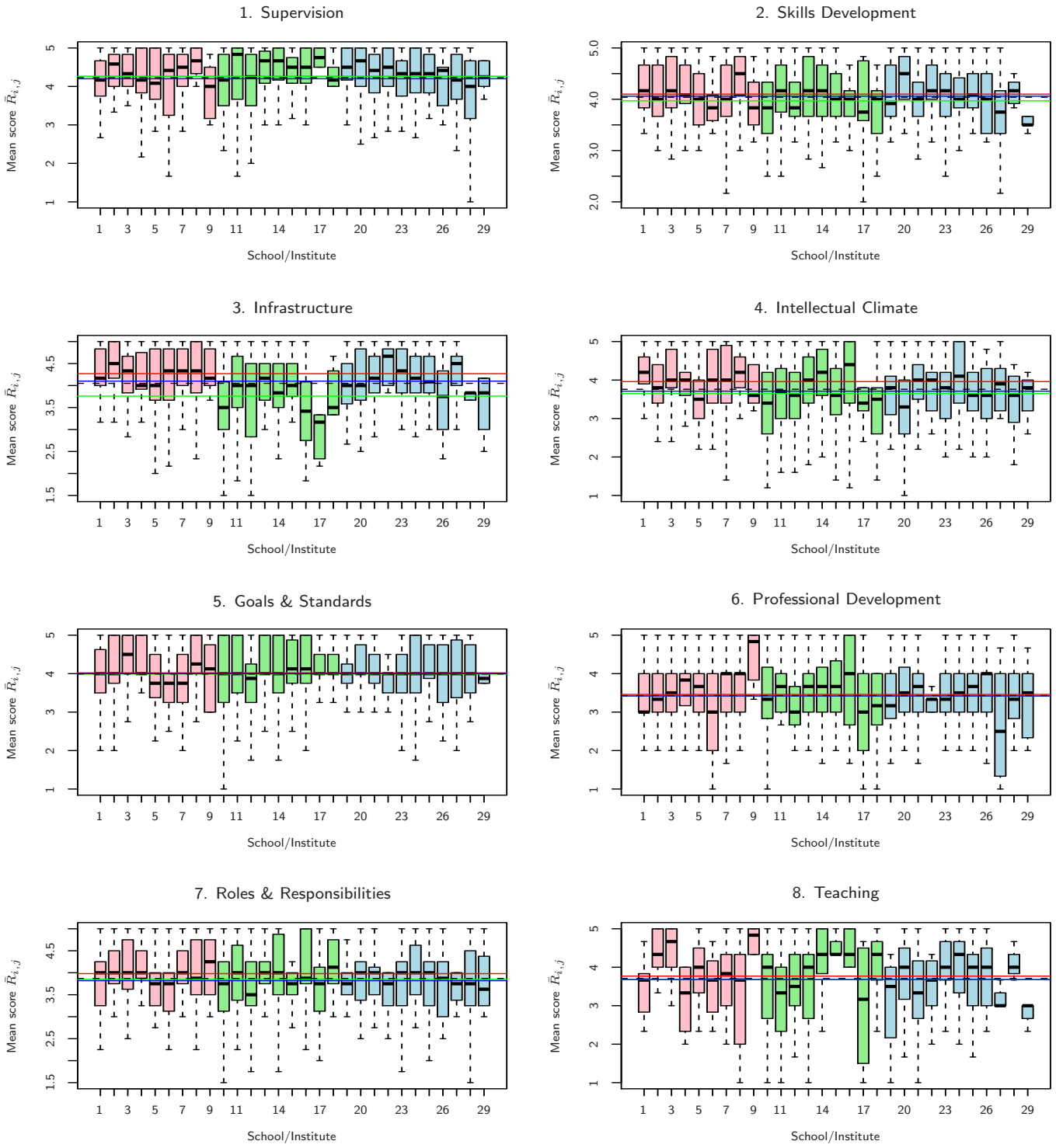

Figure 1: Boxplots of mean scores $\bar{R}_{i,j}$ for some parts of the PRES: Colours indicate Schools/Institutes belonging to FMS, HaSS and SAgE. A list of Schools/Institutes, given here by numbers 1–29, is given on page 3. The horizontal solid lines indicate faculty averages (colour–coded to match the boxplots), and the horizontal broken lines show University averages.

Skills Development: Faculty average for SAgE in line with the University average, whilst the averages for FMS and HaSS are a little higher/lower (respectively); results from ION, NICR, BIOL and NIRES seem substantially different from Faculty/University averages – unfortunately for ION and NIRES, results seem poor.

Infrastructure: Noticeable differences between all Faculty averages, with FMS coming out best, on average, and HaSS fairing worst; in HaSS, scores for ECLS, SHIS and in particular SLAW stand out as being quite low in comparison to University averages; in FMS, many Institutes perform well compared to the University average; in SAgE, CHEM seems to perform very well compared to MAST and, in particular, MECH.

Intellectual Climate: FMS average noticeably higher than both HaSS/SAgE and University averages; IAH and NICR both seem to perform well compared to the University average; conversely, IHS, SLAW and SMLS all seem to fall a bit short of the University average.

Goals and Standards: Hardly any difference between Faculty averages; ICM, NICR, SAPL, SLAW and SMLS all seem to perform quite well, whilst ION, CHEM and NIRES perform quite poorly relative to Faculty/University averages.

Professional Development: Very little difference between Faculty averages; results from CHEM seem a little low relative to these averages, and those from PSYC are notably high.

Roles & Responsibilities: FMS average noticeably higher than average from HaSS and particularly SAgE.

Teaching: Very little difference between Faculty averages, although FMS slightly higher; results generally variable between Schools/Institutes; in FMS, results from ICMB, ICM and PSYC stand out as being higher than average, whilst those from IAH are a little low; in HaSS, scores from SELS and particularly SGPS and SHIS are high; in SAgE, scores from CIVG and MECH are quite high, compared to relatively low scores from MATH and NIRES.

Key findings

Some informal comparisons of student responses to each area in the key set have been made. Figure 1 can be used to compare the *full distribution* of aggregate scores for each of these areas between different Schools/Institutes; it can also be used to compare each School/Institute with their Faculty and the University as a whole. Faculties can be compared with each other, and to the University as a whole. Tables 5–7 in the Appendix give numerical results (some of which are shown in the boxplots in Figure 1, e.g. median, quartiles). We have made some observations from these summaries above; we wait until more formal analyses have been performed before reading too much into these.

2.1.2 Other sections of the PRES

We now consider student responses to the other sections of the PRES.

Thesis examination
There were four questions in section 9 of the PRES relating to the students' experience of the thesis examination. At the time of the PRES, only 15 responders had actually taken their *viva voce* examination. Thus, comparisons between Schools/Institutes (for example) are not sensible here, and we report some general findings for all 15 students who responded.

For each student, we treat the four responses relating to thesis examination as we did for the other eight sections of the PRES previously, i.e. we find the mean response for each student across all four questions. Doing so gives an overall mean and median of 4.13 and 4.67 (respectively); the standard deviation is 1.16 and the inter–quartile range is (3.75, 4.88). Of particular interest here, which these summaries hide, might be the 20% of students who strongly disagreed with the statement "The thesis examination process was fair" (although this *does* correspond to just three students).

Expectations and importance of certain factors
Section 10 of the PRES asked students how various aspects of their degree had met with their expectations. These aspects cover areas of the PRES already summarised, e.g. "Supervision". The score range was $-3 \rightarrow +3$, where a score of zero indicates these things had met with the student's expectations; scores below zero indicate the student's experience is much more negative than their expectations, and a score above zero indicates the student's experience is much better than what they expected. Figure 2 shows how results vary by School/Institute for each of the areas students were asked about. PSYC seems to stand out as having rather low scores for students' expectations in all areas other than "Facilities". SHIS has relatively low scores for student expectations in the areas of "Research skills", "Transferable skills" and "Facilities". NIRES, on the other hand, seem to have relatively high scores for student expectations in *all* areas. Some other Schools/Institutes have scores for student expectations that seem high/low relative to their Faculty averages/University averages in certain areas, as indicated by their boxplots being substantially above/below these averages, respectively (e.g. SLAW in "Supervision" and MECH in "Research Environment"). Interesting also is that the SAgE faculty averages often exceed those from FMS and HaSS.

In section 11 of the PRES students were also asked about how important they thought the areas shown in Figure 2 are, in terms of successfully completing their research degree programme; students were then also given the opportunity to comment. There were very little differences between Schools/Institutes here. On an importance scale of 1→5, 1 being "not important at all" and 5 being "very important", the following percentage of students rated at least 4:

Supervisory support & Guidance	96.5%
Opportunities to develop research skills	91.2%
Opportunities to develop transferable skills	79.1%
Access to appropriate facilities	91.6%
The research environment	84.5%
Guidance on Institutional standards/expectations	79.4%

Student comments were rather limited; however, several students commented on their lack of understanding about exactly what "transferable skills" are – for example:

> "don't know what these are"
>
> "I don't quite understand what 'transferable skills' are."
>
> "not clear what is meant by 'transferable' "
>
> "it's just fluff"

Some comments which seem to support the graphs in Figure 2 include the following:

> "Being in this stage of my study is because of my supervisor's support and encouragement" *(Student from SLAW about supervisory guidance)*
>
> "He [supervisor] gave me full support", "Having supervisors directly related to my research", "I changed supervisor for dissertation which was extremely helpful" *(Students from SAPL about supervisory guidance)*
>
> "I think that the library is lacking a lot of key texts for PhD students in [the School of] GPS", "Language training needs are important, yet the university is not fully behind me" *(Students from SGPS about access to appropriate facilities)*
>
> "At Law School, research students have to share the computer and working space..." *(Student from SLAW about access to appropriate facilities)*
>
> "Has been old concern", "I do not like the system of hot–desking, and could work more efficiently with a dedicated space", "...students who work off campus... might be a bit isolated" *(Students from ECLS about research environment)*
>
> "Was a part–time student and did most of my writing at home so this is not important", "I work at home as a mature part time student, so this aspect is not important to me" *(IHS students about research environment)*

Personal factors

Students were given three statements in the section on "Personal factors" to rate their agreement with. These correspond to the student's support from friends and family, the student's support from their employer, and whether or not the financing of their degree places a strain on personal finances. We consider this section separate from the main set because a score of 5 on the Likert scale for the statement about personal finances corresponds to a negative outcome, rather than a positive outcome as for the other two statements, making an average calculation for personal factors, as a whole, unmeaningful. Also, we might not expect results here to vary that much between Schools/Institutes.

- Of those that responded, 88% gave a score of at least 4 to the statement "My friends and family are supportive of my research degree"; mean scores by School/Institute range from 4.3 to 5 for this question.

- Nearly half of those asked said the statement "My employer is supportive of my research degree" was not applicable to them; 82% of the remainder agreed with this with at least a score of 4, and mean scores by School/Institute range between 3.5 (CHEM) to 5 (for DClinPsy students).

- Students from FMS agreed least with the statement "the financing of my degree places a strain on my personal finances" (mean score 2.89), with students from HaSS agreeing most (mean score 3.52). In fact, there are statistically significant differences between the level of agreement with this statement for students from each of the three Faculties.

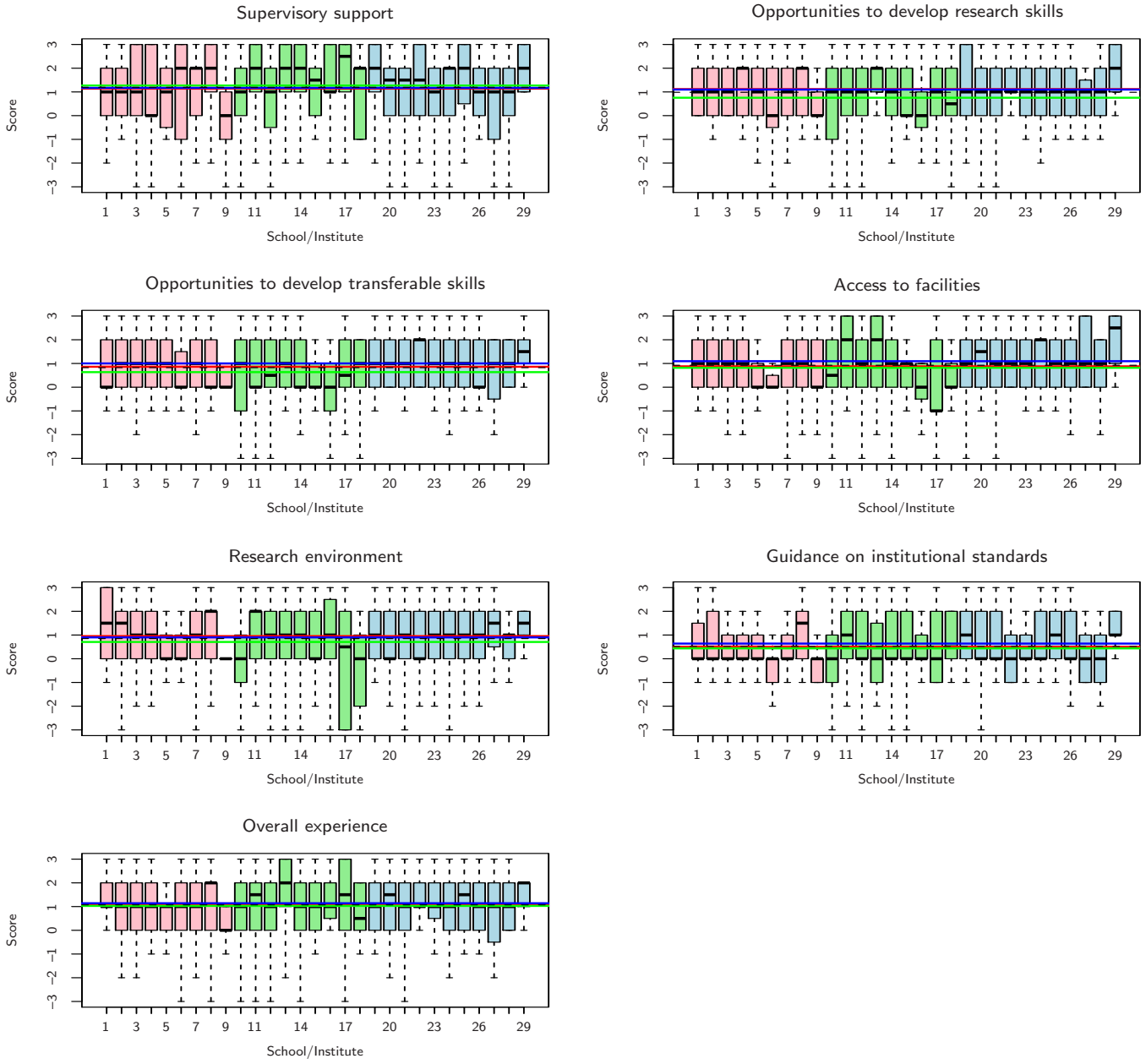

Figure 2: Boxplots of scores for expectations. Colours indicate Schools/Institutes belonging to FMS, HaSS and SAgE. A list of Schools/Institutes, given here by numbers 1–29, is given on page 3. The horizontal solid lines indicate faculty averages (colour–coded to match the boxplots), and the horizontal broken lines show University averages.

Key findings

Since only 15 responders had actually completed their *viva voce* examination at the time of the PRES, we have given simple overall summaries of student satisfaction here.

Figure 2 summarises responders' scores for "expectations", a score of zero indicating the student's expectations have been met, and positive/negative scores deviating from this accordingly. The reader can use Figure 2 to draw their own conclusions here for individual Schools/Faculties. In particular, we note that PSYC, on average, score poorly in most areas here; conversely, NIRES students' expectations have, on average, been exceeded in all areas. On average, SAgE students' expectations have been exceeded to a higher degree than their HaSS/FMS counterparts.

Generally, students think that supervisory support and guidance is the most important factor in successfully completing their research degrees; we have picked out some student comments that support this view. They find the opportunity to develop transferable skills least important; indeed, it was clear from student comments that some do not understand what "transferable skills" actually are.

It is clear that students value the support of their friends and family. It appears that FMS students feel the strain on personal finances as a result of doing a research degree least, with HaSS students feeling this most.

2.1.3 Question 15g

As discussed in section 1.3.1, question 15g in the PRES – relating to how a student's overall experience has matched with their expectations – has been suggested as an overall summary of PRES performance. The full distribution of student scores, by School/Institute, by Faculty and for the University as a whole can be seen in the last plot of Figure 2.

Key findings

Almost all Schools/Institutes have their middle 50% of scores above zero, suggesting that the expectations of the majority of students in most Schools/Institutes have at least been met. MATH stands out here as being the only School with a lower quartile below zero. The expectations of IAH, SAPL, CHEM and NIRES students seem to have been exceeded most greatly – the distribution of scores for these Schools/Institutes seems substantially above zero, and above their Faculty averages and the University average as a whole.

The following Schools/Institutes had at least 90% of their student responders scoring at least zero here: AFRD, BIOL, CHEM, CIVG, EECE, ICM, ICMB, MAST, MECH, MSGS, NIRES, SAPL and SHIS, with *all* students in CHEM, MECH and NIRES scoring at least zero.

2.2 More formal analyses

The comparisons made in section 2.1 are informative, although rather subjective. We now proceed to more formal comparisons; specifically, in section 2.2.1 we look for *statistically significant* differences between mean scores for each School/Institute and corresponding Faculty/University averages, and in section 2.2.2 we compare Schools/Institutes with each other (within each of the three Faculties separately).

2.2.1 School/Institute comparisons with Faculty and University

Figure 3 shows, for each School/Institute, the difference between mean scores for that School or Institute and the mean for the Faculty to which that School/Institute belongs. A positive difference indicates that a School/Institute has a mean score greater than its Faculty's mean score; a negative difference indicates that the School/Institutes mean score is lower than its Faculty's average. Schools/Institutes have been labelled according to the key on page 3. But how do we know if these differences are *statistically significant*?

Statistical details

For each set of School–Faculty and School–University mean differences, we use a t–test with the following (respective) null hypotheses:

H_0 : There is no difference between the School mean and its Faculty mean

H_0 : There is no difference between the School mean and the University mean

Here, and throughout the rest of this report, we assume a 5% level of significance – i.e. we reject H_0 in a particular comparison if the associated p–value < 0.05.

Symbols have been attached to the bars in Figure 3 in cases where the difference between School/Institute and Faculty means is statistically significant (see box above): $\triangle$ indicates that a School/Institute mean score is significantly *higher* than its Faculty average, whilst $\bigtriangledown$ indicates a significantly *lower* than Faculty average mean score. Figure 4 shows exactly the same information, but now for differences between School/Institute mean scores and the scores for the University as a whole. Table 1 summarises the results from Figures 3 and 4; notice that conclusions drawn here lend support to some of the more informal comments made in section 2.1.1.

	Significantly **higher** or **lower** than	
	Faculty average	University average
Supervision	NICR	ICMB, NICR
Skills Development	NICR, ION BIOL	IAH, ICM, NICR, ION ECLS BIOL
Infrastructure	ICMB SAPL, ECLS CHEM	ICMB, ICM, MSGS ECLS, SACS, SELS, SHIS CHEM
Intellectual Climate	IAH, IHS SELS, ECLS COMP	IAH, ICM, MSGS ECLS
Goals & Standards	ICM, ION, MSGS	ICM, ION
Professional Development	ION SACS	ION SACS
Roles & Responsibilities	ION ECLS	ICM ECLS
Teaching Opportunities	ICM SELS, SHIS CIVG	ICMB, ICM SELS, SHIS

Table 1: Schools/Institutes whose mean scores were significantly higher or lower than (i) their Faculty average and (ii) the University average, as shown in Figures 3 and 4. Those Schools/Institutes not mentioned have mean scores that are *not* significantly different from their Faculty mean/University mean.

Key findings

Figures 3 and 4 compare (respectively) average scores for each School/Institute with (i) their Faculty's average scores and (ii) the University averages as a whole; comparisons are made – statistically – for each area in the key set. Table 1 then highlights Schools/Institutes that are performing significantly above, and below, both their Faculty average and the University average, in each of these areas. Some of the earlier (more informal) observations made in section 2.1.1 are now supported with *statistical significance*. Although it should be clear from Figures 3 and 4 which Schools/Institutes are performing above/below average, we pull out some of the more convincing observations now.

ICM seems to do well, with the highest number of positive (significant) departures from its Faculty averages/the University averages as a whole. Generally, the Faculty of Medical Sciences have a much higher number of (significant) positive departures from the University's averages than the other two Faculties. In contrast, ION scores are, on average, below FMS/University scores in a few areas; similarly, ECLS average scores are also significantly below the HaSS/University averages in a number of areas. There could also be a problem with "Infrastructure" in HaSS, with four Schools having scores here that are significantly lower than the University's average.

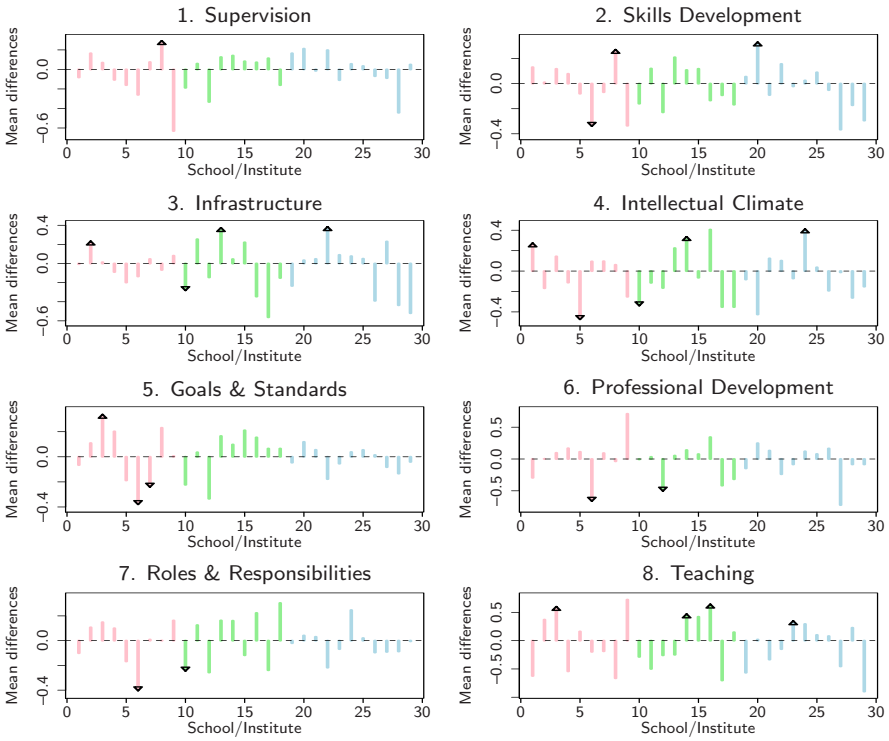

Figure 3: Mean differences from Faculty averages. △ indicates that the School/Institute mean is significantly *higher* than the Faculty mean; ▽ indicates that the School/Institute mean is significantly *lower* than the Faculty mean. Colours indicate Schools/Institutes belonging to **FMS**, **HaSS** and **SAgE**.

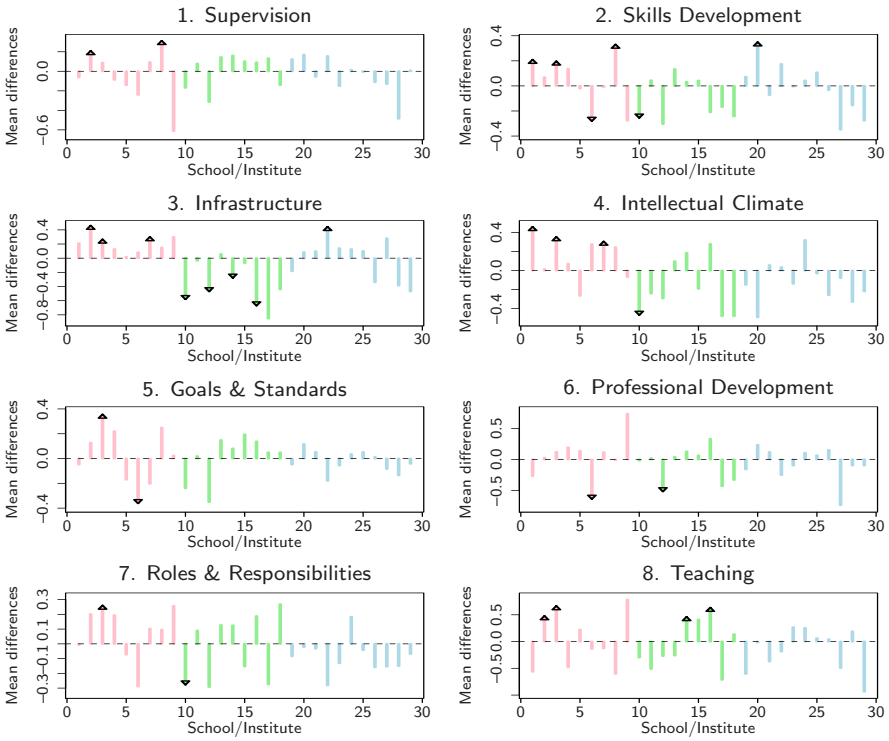

Figure 4: Mean differences from University averages. △ indicates that the School/Institute mean is significantly *higher* than the University mean; ▽ indicates that the School/Institute mean is significantly *lower* than the University mean. Colours indicate Schools/Institutes belonging to **FMS**, **HaSS** and **SAgE**.

2.2.2 School/Institute comparisons with each other

In section 2.2.1 we compared mean scores for each School/Institute with their Faculty's mean, and with the overall University mean, for questions in the main set. We now compare Schools/Institutes from the same Faculties *with each other*.

Statistical details

We use Analysis of Variance (ANOVA) to perform within–Faculty comparisons between Schools or Institutes. Within each Faculty, and for each section of the key set, we test the null hypothesis

$$H_0 \quad : \quad \text{School/Institute averages are equal.}$$

We also compare Faculties with each other in the same way, testing

$$H_0 \quad : \quad \text{Faculty averages are equal.}$$

When a p–value for an ANOVA suggests there is evidence to reject H_0, we use *Tukey's multiple comparisons* procedure to detect exactly where the differences lie (i.e. between which pairs of Schools/Institutes or between which pairs of Faculties there are significant differences).

The assumption of Normality in ANOVA is plausible here (as it is for the t–tests in section 2.2.1). For example, for each of the eight sections in the key set of the PRES, we have summarised results for every student responder by their mean score for the set of associated questions; the *central limit theorem* tells us that our means – at least in the limit – should be Normally distributed. This, along with other assumptions implicit in ANOVA, can also be checked post–analysis by examination of model residuals. We have done this, and there appear to be no issues with these underpinning assumptions.

Table 2 below summarises our results on each of the eight sections in the main set of the PRES. We have compared mean scores for each of these areas between Schools/Institutes within each Faculty separately (so Schools/Institutes from different Faculties are *not* compared); we have also compared mean scores between the three Faculties themselves. Where significant differences *have* been detected between means of $\bar{R}_{i,j}$, a follow–up procedure has been employed to detect where these differences lie; these are reported in Table 2 also. For example, we have detected significant differences between School scores for "Infrastructure" in SAgE; the follow–up procedure has then detected that these differences are between (Chemistry and Marine Science & Technology) and (Chemistry and Mechanical & Systems Engineering). Where significant differences have been detected, the School/Institute given first in the parentheses is that with the highest scores of the pair. So, in the previously quoted example, "Infrastructure" scores for Chemistry appear to be significantly higher than those for both Marine Science & Technology and Mechanical & Systems Engineering.

Notice that, in these more formal comparisons, we have detected differences that may not have been immediately apparent from the boxplots in Figure 1; these comparisons have, however, also served to confirm some of the points of interest made in section 2.1.1 (e.g. significant differences between Faculty averages for "Infrastructure").

PRES Section	**FMS**		**HaSS**		**SAgE**		**University**	
	p–value	Differences	p–value	Differences	p–value	Differences	p–value	Differences
Supervision	0.026	(NICR,PSYC)	0.156	—	0.540	—	0.575	—
Skills Dev.	0.023	(NICR,ION)	0.131	—	0.156	—	0.042	(FMS,HaSS)
Infra–structure	0.345	—	0.015	(SAPL,ECLS)	0.040	(CHEM,MAST) (CHEM,MECH)	0.000	(FMS,SAgE) (FMS,HaSS) (SAgE,HaSS)
Intell. Climate	0.016	(IAH,IHS) (ICM,IHS) (MSGS,IHS)	0.007	(SELS,ECLS)	0.205	—	0.000	(FMS,HaSS) (FMS,SAgE)
Goals & Stand.	0.002	(ICM,ION) (ICM,MSGS)	0.133	—	0.990		0.876	—
Prof. Dev.	0.031	(PSYC,ION)	0.264	—	0.238	—	0.877	—
Roles & Resp.	0.173	—	0.059	—	0.882	—	0.039	(FMS,SAgE)
Teaching Opp.	0.001	(ICM,IAH) (ICM,NICR)	0.020	(SGPS,NUBS)	0.047	(CIVG,CEAM) (CIVG,AFRD)	0.701	—

Table 2: ANOVA results for within Faculty School/Institute comparisons and Faculty comparisons.

Key findings

The Faculty of Medical Sciences has significant differences between some of its Institutes in all but two areas of the key set. In half of pairs which are significantly different to each other, ICM is the best performer (we might expect this given the results in section 2.2.1). IHS and ION are often the worst performers in pairs where differences have been detected. The high–flyers in SAgE are CIVG and CHEM (Teaching Opportunities and Infrastructure, respectively). More generally, where significant differences between Faculties have been detected, FMS appear strong relative to the other two Faculties. It might be possible for relatively under–performing Schools/Institutes (the second entries in the parentheses) to seek advice from those that are performing well (the first entries in the corresponding parentheses).

3 Newcastle PRES: Trends over time

The aim of this section is to compare results from the Newcastle PRES 2011 with Newcastle PRES results from previous years. The PRES survey at Newcastle was also carried out in 2008 and 2009 (not in 2010); we will compare Schools'/Institutes', Faculties' and the University's scores from each section in the key set in the PRES 2011, as well as question 15g, to those from previous years. Again, we will make use of simple graphical summaries as well as more detailed statistical analyses.

Statistical details

Aggregated scores for each student, for each area in the key set, are considered as observations from *response variables* Y_k, $k = 1, \ldots, 8$, where k is an indicator of the section in the key set (e.g. $k = 1$ corresponds to "Supervision", $k = 2$ corresponds to "Skills Development", etc.). For each Y_k we have 2,551 observations; 919 from PRES 2011, 825 from PRES 2009 and 807 from PRES 2008. Each observation on Y_k has an associated time indicator, t_k, where $t_k = 1$, $t_k = 2$ or $t_k = 3$ for observations from 2008, 2009 or 2011 (respectively). We perform simple linear regressions of Y_k on t_k; i.e. we fit the model

$$Y_k = \beta_0 + \beta_1 t_k + \epsilon_k, \qquad \epsilon_k \sim N(0, \sigma^2),$$

separately for each section of the key set; in each regression, a significant slope (i.e. a p–value of less than 0.05 for the estimated slope $\hat{\beta}_1$) will suggest that there is a significant (linear) trend in student scores over time. We also investigate the significance of trends through time for Faculties and the University as a whole. Confidence bounds for our estimated trends can also be obtained. This procedure can also be performed for responses to question 15g.

For illustration, Figure 5 shows a plot of scores for "Supervision" (averaged across all questions in the section on "Supervision" in the PRES) for students in NICR, in the 2008, 2009 and 2011 PRES. The asterisks indicate overall means for NICR in each year, and these have been joined to show the average transition year–on–year (red line). The dashed line is the fitted trend (linear regression) line (see the box above); the slope of this line is statistically significant. The outer dotted lines correspond to the 95% confidence limits for the fitted trend. This plot shows that scores for "Supervision" in NICR have improved significantly year–on–year. Figure 6 shows similar plots for all other sections of the key set; plots are only shown for those Schools/Institutes with significant trends, and the fitted trend lines (with confidence intervals) have been omitted to avoid having overcrowded plots. Table 3 reports those Schools/Institutes/Faculties that have seen significant changes in mean PRES scores over time; this table also reports any trends at the University level.

It should be noted here that it was not possible to estimate possible trends for some Schools or Institutes. For example, in PRES 2008, Biology and Psychology were part of the same School, and various groups that existed in the Faculty of Medical Sciences in 2008 were restructured before PRES 2009/2011. Thus, we investigate trends over time only for those Schools/Institutes that have been in existence continually since 2008.

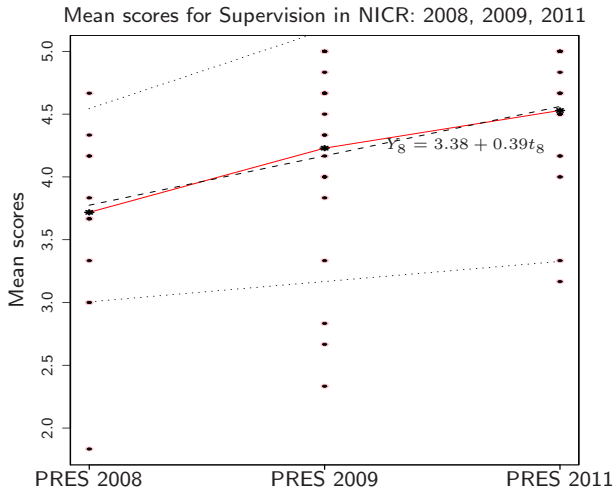

Figure 5: Plot of mean scores in NICR for Supervision – PRES 2008, 2009 and 2011. The points show each student's mean score across all questions relating to Supervision (8 students in 2008, 12 students in 2009 and 8 students in 2010). The solid red line joins the overall mean scores in NICR for Supervision in each year. The dashed line is the fitted trend line (which is statistically significant with a p–value of 0.002); the outer dotted lines are the associated lower and upper 95% confidence bounds.

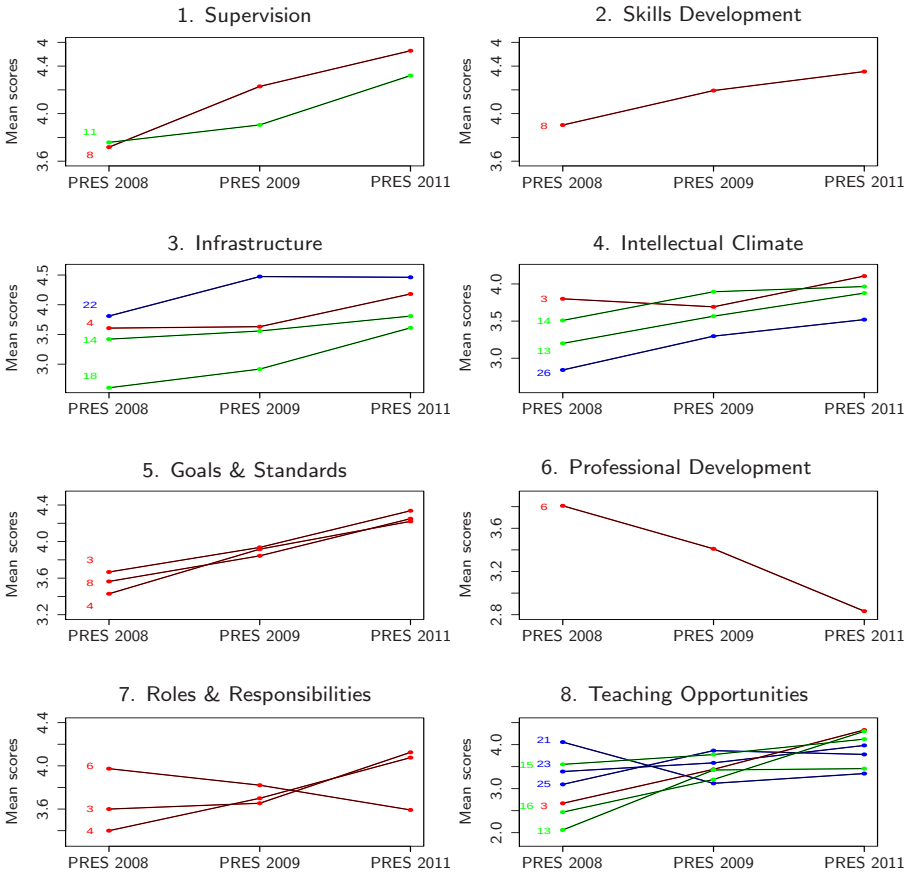

Figure 6: Plots showing the transition of mean scores, by School/Institute, for each component of the key set. Traces are only shown for Schools/Institutes whose fitted trends are statistically significant; the School/Institutes are given down the left–hand–side of each trace, using the key on page 3.

	Significant change over time?	
	Positive change	Negative change
Supervision	NICR, NUBS HaSS University	
Skills Development	NICR	
Infrastructure	IHG, SELS, SMLS, CHEM FMS, HaSS University	
Intellectual Climate	ICM, SAPL, SELS, MAST FMS, HaSS, SAgE University	
Goals & Standards	ICM, IHG, NICR FMS, SAgE University	
Professional Development	HaSS	ION
Roles & Responsibilities	ICM, IHG FMS, HaSS University	ION
Teaching Opportunities	ICM, SAPL, SELS, SHIS, SGPS, CIVG, EECE FMS, HaSS University	CEAM

Table 3: Schools/Institutes showing a significant (linear) trend (Positive/Negative) in mean scores from the Newcastle PRES in 2008, 2009 and 2011. Also indicated are trends over time in results by Faculty, and for the University as a whole.

Key findings

Table 3 summarises our findings here. We see that generally, the University as a whole has shown significant improvement, year–on–year, in all sections of the key set of the PRES other than "Skills Development" and "Professional Development". At a Faculty level, HaSS have improved significantly in all areas of the key set other than "Skills Development" and "Goals and Standards"; FMS have shown significant improvement in all areas other than "Supervision", "Skills Development" and "Professional Development"; the only areas SAgE have shown significant improvement in are "Intellectual Climate" and "Goals and Standards". However, it should be noted that neither the University, nor each of the Faculties individually, showed any significant *deterioration* across each of the areas in the key set. Standing out at the School/Institute level are NICR, ICM and SELS, showing significant improvement in at least three areas of the key set; conversely, ION have shown significant deterioration in "Professional Development" and "Roles and Responsibilities", and CEAM have shown deterioration as far as "Teaching Opportunities" are concerned.

When comparing year–on–year percentages of students scoring at least zero in question 15g, we find that there are no significant trends for *any* Schools/Institutes (results for question 15g are not shown here).

4 Comparing Newcastle PRES 2011 to other groups

In this section we compare results from the PRES 2011 at Newcastle to (i) the national picture, and (ii) to results from other Russell Group institutions also taking part in the PRES in 2011. We do not have results for other institutions at an individual student level; rather, for each question in the PRES we have the proportion of responders scoring 1, 2, 3, 4 or 5. As already discussed, for most questions (all in the key set) the higher the score, the more positive the response. Thus we consider, for each question in the key set of the PRES, the proportion of students scoring at least 4. For question 15g we consider the proportion of students scoring at least zero (recall that, for this and other questions in section 10 of the questionnaire, a scale from $-3 \rightarrow +3$ was used). Note that, because we are not working at a School/Institute level here, but are interested in results for Newcastle as a whole, it is feasible to compare Newcastle to other institutions for *all* questions in the key set. For example, there were six questions relating to "Supervision"; we now consider each of these individually instead of aggregating across all six. Similarly, there were 6, 6, 5, 4, 3, 4 and 3 questions for "Skills Development", "Infrastructure", "Intellectual Climate", "Goals & Standards", "Professional Development", "Roles & Responsibilities" and "Teaching Opportunities", respectively, giving 37 questions in total for the key set.

Statistical details

To compare the proportion of Newcastle students scoring at least 4 out 5 (or at least zero for question 15g) to that from all other institutions taking part in the PRES 2011, we can use a *test of equal proportions*. Our null hypothesis is

$$H_0 \quad : \quad \text{Newcastle proportion} = \text{proportion from all other Institutions},$$

when we compare Newcastle to all other institutions; comparing Newcastle to other Russell Group institutions taking part in PRES, we have

$$H_0 \quad : \quad \text{Newcastle proportion} = \text{proportion from other Russell Group institutions}.$$

As before, we work at the 5% level of significance, and we test the above hypotheses for all 37 questions in the key set, as well as question 15g.

As previously reported, there were 926 responders to the PRES at Newcastle; from *all* other institutions there were another 30,276 responders, and from all other Russell Group institutions there were 13,968 responders. Figure 7 shows plots of the percentage of Newcastle students scoring at least 4 (or at least zero for question 15g) against the corresponding percentages for (i) all other institutions in PRES 2011 (top), and (ii) all other Russell Group institutions in PRES 2011 (bottom). There are 38 points in each plot in total – one for each question in the key set (37 questions), plus question 15g. These points are numbered to correspond to their section in the PRES – **1** for "Supervision", **2** for "Skills Development" and so on (see section 1.2); the point labelled "S" is for question 15g (used as an overall **S**ummary of PRES). The diagonal line represents the 'line of equality' – i.e. if a point lies on this line then the percentage of students scoring at least 4 is exactly the same for both groups of students being considered; points above this line mean a higher percentage for Newcastle, points below the line mean the other group has a higher percentage. Points marked with an asterisk correspond to comparisons giving a statistically significant difference (see the box above).

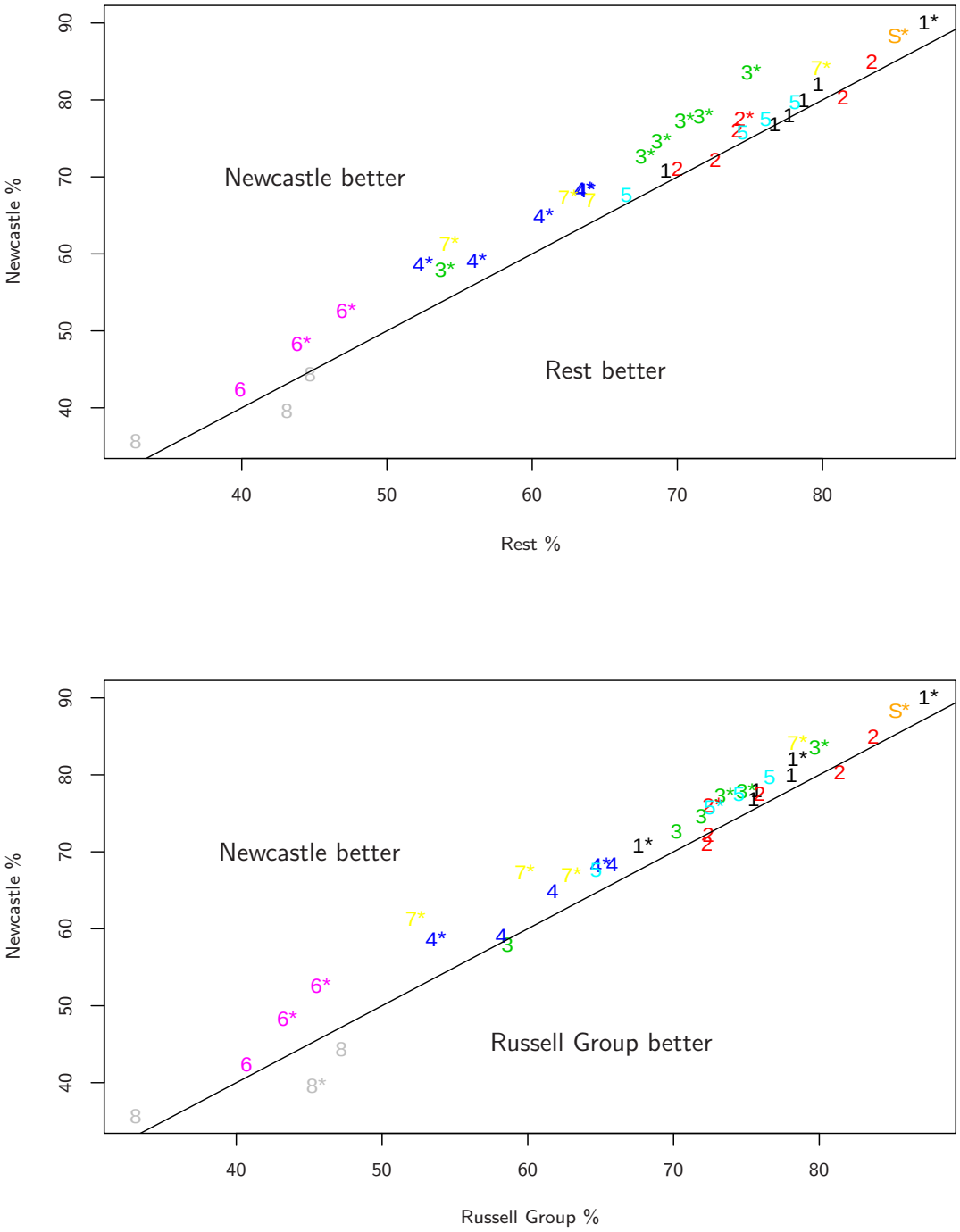

Figure 7: Percentage of Newcastle students scoring at least 4 out of 5 for questions in the key set (0 out of 3 for question 15g) against corresponding percentages for (i) all other institutions in PRES 2011 (top) and (ii) all other Russell Group institutions in PRES 2011 (bottom). Numbers correspond to areas of the key set (see section 1.2); "S" corresponds to question 15g, used as an overall summary of PRES. Points marked with an asterisk denote statistically significant differences. The diagonal line is the line of equality.

Key findings

It is clear from Figure 7 that Newcastle fairs better than the rest of the institutions taking part in PRES 2011, as a whole, in almost all questions in the key set *and* question 15g. This is with the exception of 2 questions relating to "Teaching Opportunities" and 2 questions relating to "Skills Development". Many of the questions in which Newcastle gains a higher percentage are shown to be statistically significantly higher; most notably, Newcastle percentages are significantly higher for *all* questions relating to "Infrastructure" and "Intellectual Climate", as well as question 15g. In the few cases where the "rest" fair better than Newcastle, these differences are not significant.

Comparing Newcastle to the other Russell Group institutions, again the picture is bright: Newcastle fairs better in 32 out of the 38 questions in total. There are fewer significantly higher scores for Newcastle this time, although Newcastle scores are significantly higher for *all* questions in "Roles & Responsibilities". The other Russell Group institutions perform significantly better than Newcastle for one question related to "Teaching Opportunities". Once again, Newcastle's percentage for question 15g is significantly higher than the other group considered as a whole.

5 Statistical analysis for survey non–response

It is well–known that differential response rates for different categories of respondents can induce bias in results, which can then distort any post–survey between–groups comparisons (as in sections 2 and 3). One way to identify and account for potential *response bias* is through the use of an *inverse probability of response rate* to adjust survey responses. A toy example is now given to illustrate.

Suppose we are interested in responses to a particular survey question. In this question, the responder is asked to rate, on a scale of 1 to 5, their agreement to a given statement. Suppose further that responders can be categorised into three groups: A, B or C (perhaps age groups, or groups of people with some attribute A, B or C). Suppose there are three people in each of groups A, B and C; their responses to this question are given in the first row of the table below (labelled "Full response"). Notice that the mean response across all groups is

$$\frac{2+2+2+3+3+3+4+4+4}{9} = 3.$$

Group	**A**			**B**			**C**		
Full response	2	2	2	3	3	3	4	4	4
Partial response	2	2	2	3	?	3	?	?	4
Response Probability	1			$\frac{2}{3}$			$\frac{1}{3}$		
Weights	1			$\frac{3}{2} = 1.5$			$\frac{3}{1} = 3$		

Now suppose everyone in group A responded, but only two people in group B responded and only one person in group C responded. The second row of the table above gives this "Partial response". The mean response is now

$$\frac{2+2+2+3+3+4}{6} = 2.667.$$

We know that this mean is lower than it should be, because we know what the true mean is (3, from the "Full response"). Even without this information, we might suspect our mean from the "Partial response" is biased. It seems that people from group C tend to give higher ratings and two–thirds of this group did not respond; people from group A give low ratings, and they *all* responded! Thus, we might speculate that the mean should be higher than 2.667 to account for the 'high–scorers' who did not respond. How can we do this?

The third row in the above table gives the associated *response probability* for each group – just the proportion responding. The fourth row then gives the *inverse probability weights* – just the inverse of the probabilities shown in the row above. We now weight each response we observe (x) by the values shown in the fourth row (w), to obtain a weighted mean which corrects for bias in non–response, i.e.

$$\frac{\sum wx}{\sum w} = \frac{(2\times 1)+(2\times 1)+(2\times 1)+(3\times 1.5)+(3\times 1.5)+(4\times 3)}{1+1+1+1.5+1.5+3} = \frac{27}{9} = 3, \qquad (1)$$

which has been adjusted upwards to account for the 'large' number of non–responders from group C (recall that those from group C seem to rate more highly). In fact, in this example, we have recovered the true mean!

5.1 Estimating the response probabilities

5.1.1 Dependence on covariates: the logistic regression equation

What if we're interested in the effects of several *covariates* (e.g. age, gender etc.) on the response rate within each School/Institute? Then the way we calculate the response probabilities will have to change. It could be that such covariates influence the response probability, and so each student invited to complete the PRES will have a response probability depending on their particular values/levels of these covariates. The standard way of modelling dependencies in statistics is through regression; for *binary* outcomes (e.g. student i responds/does not respond), *logistic regression* is often used – this will build into the response probabilities any dependence on covariates.

Statistical details

We let the *response* variable $y_{i,j}$ indicate whether or not a postgraduate research student completed the PRES:

$$y_{i,j} = \begin{cases} 1 & \text{if student } i \text{ in School } j \text{ completes the survey;} \\ 0 & \text{if student } i \text{ in School } j \text{ does } \textit{not} \text{ complete the survey.} \end{cases}$$

We consider several *predictor* variables that could influence the probability that a student completes the PRES – $\Pr(y_{i,j} = 1)$ – namely: age, gender, whether the student is a home/EU/Non–EU student, and whether the student is full or part time. Then, for each School/Unit j, we use

$$\Pr(y_{i,j} = 1 | x_{1_{i,j}}, x_{2_{i,j}}, \ldots) = \frac{\exp\left\{\beta_0 + \beta_1 x_{1_{i,j}} + \beta_2 x_{2_{i,j}} + \ldots\right\}}{1 + \exp\left\{\beta_0 + \beta_1 x_{1_{i,j}} + \beta_2 x_{2_{i,j}} + \ldots\right\}}, \qquad (2)$$

where $x_1, x_2, \ldots$ are covariates representing age, sex, student origin (e.g. home/EU/non–EU) and study status (full time/part time). Equation (2) is known as the *logistic regression equation*, the form of which ensures that the outcome will always lie in the range $(0, 1)$ for probabilities, whilst at the same time allows for a dependence on covariates.

5.1.2 Application to the PRES data

To obtain response probabilities via logistic regression (see box above) for the PRES, we require, for every student invited to complete the survey, the following:

Student (i)	School/Institute (j)	Completed? ($y_{i,j}$)	Age	Gender	FT/PT	Home/EU/Non–EU
1	BIOL	0	24	Male	FT	Home
2	MATH	1	33	Female	FT	Non–EU
3	SGPS	1	23	Female	PT	Home
⋮	⋮	⋮	⋮	⋮	⋮	⋮
2137	IAH	1	29	Male	FT	EU

The PRES itself gives this information for responders; Jonathan Gray, an administrator from Student & Academic Services, provided the same information for those students who did *not* complete the PRES.

Statistical details

The variables 'Gender', 'FT/PT' and 'Home/EU/Non–EU' are all known as *factorial variables* – i.e. their outcomes are factors taking a finite number of levels. Gender, for example, has 2 levels (Male/Female). We incorporate these into our regression analysis by using indicator variables which take only the values 0 and 1. Generally, a factorial variable with k levels will require $k-1$ indicator variables; thus, 'Gender', 'FT/PT' and 'Home/EU/Non–EU' will each require 1, 1 and 2 indicator variables respectively. Defining our first predictor variable as x_1: Age, we then define another four predictor variables:

	x_2		x_3		x_4	x_5
Male	1	FT	1	Home	1	0
Female	0	PT	0	EU	0	1
				Non–EU	0	0

We can fit the logistic regression model given by Equation (2), initially using all five predictor variables $x_1, \ldots x_5$, using the statistical package **R**. A commonly–used procedure known as *backwards elimination* is then used to reduce the number of predictors until all remaining predictors are statistically significant. In fact, for our data, variables associated with Home/EU/Non–EU (i.e. x_4 and x_5) were found to be insignificant. This implies that whether a student is from the UK, the EU or outside the EU has no influence over that student's likelihood of completing the PRES. Below is some output from the package **R**, showing estimated coefficients and associated p–values (all now less than 0.05), for the remaining predictor variables:

```
                 Estimate  Std.  Error  z value  Pr(>|z|)
(Intercept)     -0.160127     0.251569   -0.637   0.52444
x1 (Age)        -0.017778     0.005889   -3.019   0.00254  **
x2 (Male)       -0.413392     0.090168   -4.585  4.55e-06  ***
x3 (Full time)   0.843965     0.129059    6.539  6.18e-11  ***
```

For our data, this gives the following logistic regression equation (on substitution into (2)):

$$\Pr(y_{i,j}=1|x_{1_{i,j}},\ldots) = \frac{\exp\left\{-0.160-0.018x_{1_{i,j}}-0.413x_{2_{i,j}}+0.844x_{3_{i,j}}\right\}}{1+\exp\left\{-0.160-0.018x_{1_{i,j}}-0.413x_{2_{i,j}}+0.844x_{3_{i,j}}\right\}}. \quad (3)$$

We find that the probability of completing the PRES depends significantly on the Age of the student, whether or not the student is male and whether or not the student is studying full time, but *not* on whether the student is a home/EU/Non–EU student. The plots shown in Figure 8 overleaf show how this probability changes with age, and according to different combinations of sex and FT/PT status. Some important observations from Figure 8:

- The older a student, the less likely they are to complete the PRES;
- No matter whether a student is full time or part time, females are more likely to respond to the PRES than males (significantly so for full time students younger than about 45);
- For both males and females, full time students are significantly more likely to complete the PRES than part time students.

We can now use Equation (3) on each student invited to complete the PRES, replacing x_1, x_2 and x_3 with that responder's age, sex and FT/PT status. Doing so gives the histograms shown in Figure 9. We clearly have bimodal distributions in both cases, corresponding to the significant differences in probabilities between FT and PT students for both sexes (as seen in Figure 8).

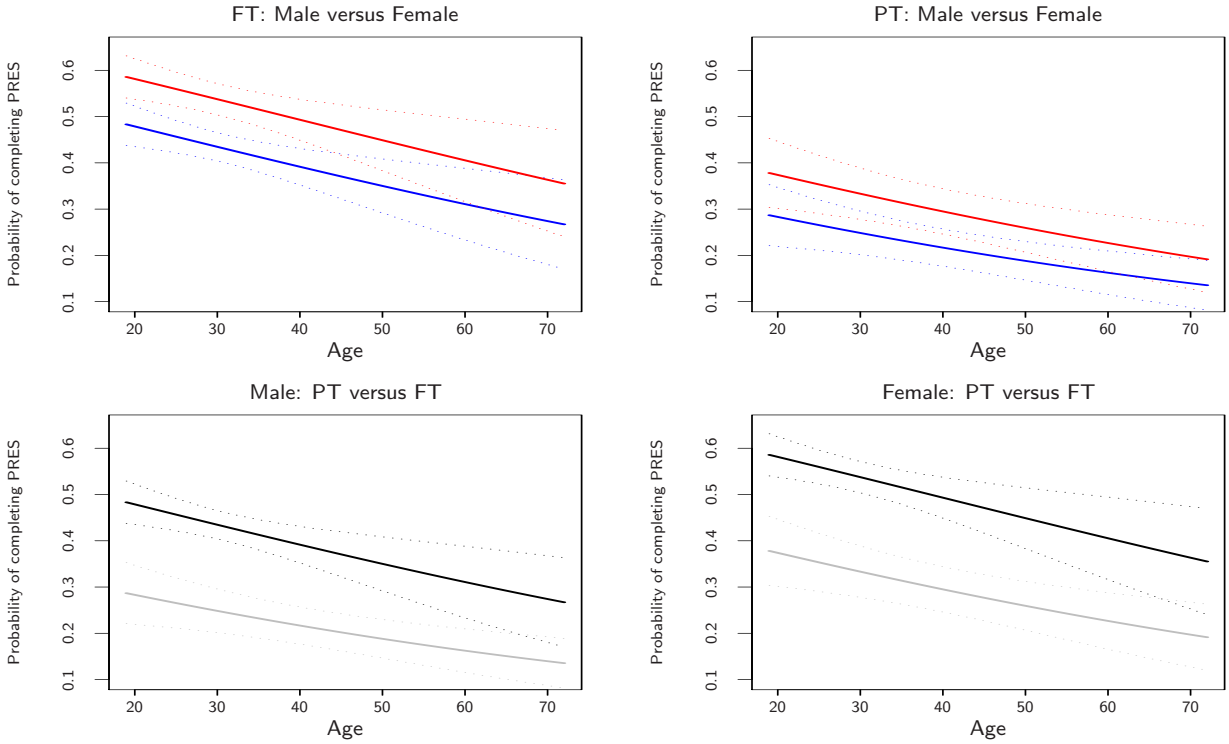

Figure 8: Probability of completing PRES versus age of student. Top left: full time students (**male** and **female**); top right: part time students (**male** and **female**); bottom left: Male students (**full time students** and **part time students**); bottom right: Female students (**full time students** and **part time students**). Dotted lines are the corresponding 95% confidence intervals for these probabilities.

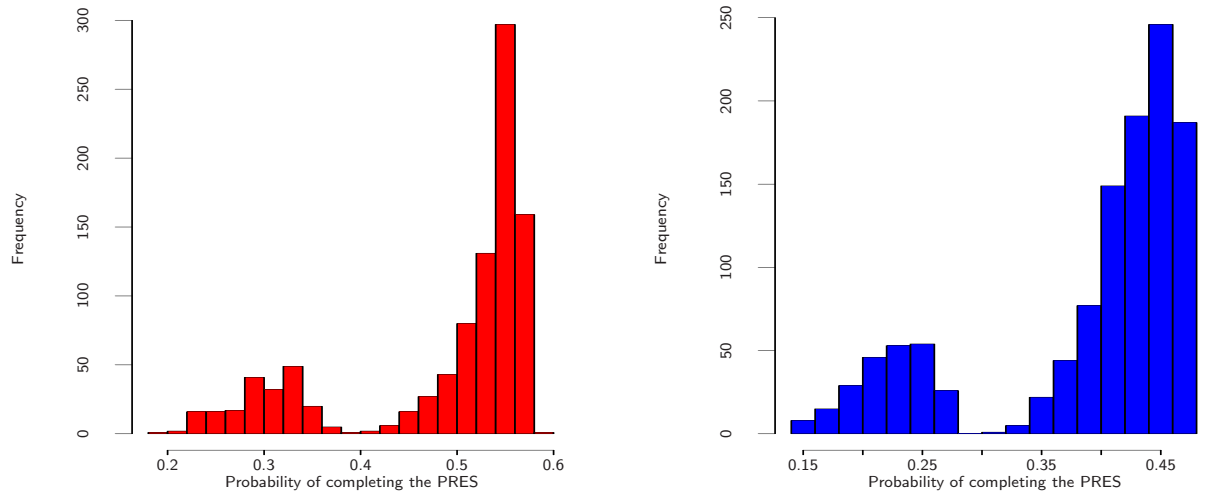

Figure 9: Histograms of estimated probabilities of completing the PRES for all students invited to complete the PRES (left: **females**, right: **males**).

5.2 Using the logistic response probabilities

We now use the logistic response probabilities obtained in section 5.1.2 (for those who actually completed the PRES) to weight scores to adjust for survey non–response.

Recall from section 1.3.2 that students were asked a series of questions within each of the eight sections of the key set; an overall summary of each section, for every responder, was then obtained by finding their mean score across each series of questions ($\bar{R}_{i,j}$ for student i in School/Institute j). We then summarised these mean scores by University, Faculty and School/Institute. Recall also that simple metrics such as means, medians and quartiles were used to compare Schools/Institutes – and Faculties – with each other, as well as with the University as a whole; we also used more formal comparisons to check for statistical significance (section 2.2).

Similar to the toy example on page 21, we calculate weights to adjust for student non–response, but now we calculate these weights from a response probability that depends on age, gender and FT/PT study status. For each student i in School/Institute j, the weight is found as $w_{i,j} = 1/\Pr(y_{i,j} = 1)$, where this probability is obtained via Equation (3). Averages for $w_{i,j}$ for each School/Institute are shown in Figure 10 below. As we would expect, the higher the average response probability for a particular School/Institute, the lower the corresponding re–scaled weight necessary.

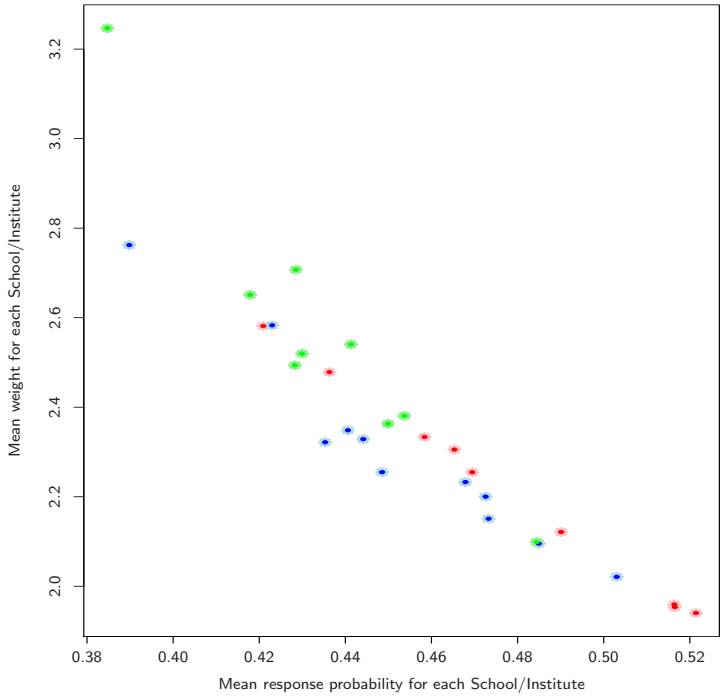

Figure 10: Plot of mean re–scaled weights against mean response probability for each School/Institute belonging to FMS, HaSS and SAgE.

The weighted mean score for School/Institute j is then obtained in the same way we obtained the weighted mean in the toy example (Equation (1) on page 21), i.e.

$$\bar{R}_{\text{adj},j} \quad = \quad \frac{\sum w_{i,j}\bar{R}_{i,j}}{\sum w_{i,j}}, \tag{4}$$

where we sum across all students i within School/Institute j. We can calculate this adjusted mean score for each category of the main set (e.g. "Supervision", "Skills Development", ...) and, as before, compare these means with Faculty/University means (also adjusted).

Figures 11 and 12 serve to reproduce the comparisons shown in Figures 3 and 4 (respectively), but now show mean differences based on adjustment for non–response. Thus, comparisons shown in Figures 11 and 12 have taken into account the varying response rates of individual students within each School/Institute (and Faculty/University as a whole) depending on age, gender and FT/PT study status. These comparisons thus avoid any bias due to non–response that we might see in section 2.2. Table 4 is equivalent to Table 1, but again for the adjusted mean scores.

Performing ANOVA to check for (i) differences between Schools/Institutes within Faculties, and (ii) differences between Faculties (as in section 2.2.2), but now using the adjusted results, reveals: significant differences between Institutes within FMS for all eight sections of the PRES covered here; significant differences between Schools within HaSS for these eight sections of the PRES; significant differences between Schools within SAgE for all but "Goals & Standards"; and significant differences between the three Faculties for "Skills Development", "Infrastructure", "Intellectual Climate" and "Teaching Opportunities". *NB: In the analyses of adjusted results, weighted sums of squares have been obtained in a similar fashion to the weighted means found via Equation (4).*

	Significantly **higher** or **lower** than	
	Faculty average	University average
Supervision	ICMB, IAH NUBS, SAPL, SELS, ECLS MAST	ICMB, IAH NUBS, SAPL, SELS, ECLS AFRD, MAST
Skills Development	ICM, IHS SAPL, SGPS, SLAW EECE	ICM, MSGS, NICR SAPL, SELS, SHIS, SLAW EECE
Infrastructure	IAH, ICMB, ICM, IHS SAPL, SGPS, ECLS, SLAW CEAM, CHEM, CIVG, MECH, NIRES	IAH, ICMB, ICM, ION, MSGS, PSYC ECLS, NUBS, SELS, SHIS, SLAW CEAM, CHEM, CIVG, EECE
Intellectual Climate	IAH, IHS SAPL, SELS, ECLS, NUBS CEAM, CHEM, COMP, BIOL, MAST	IAH, ICMB, ICM, MSGS, NICR SAPL, SELS, ECLS, NUBS BIOL, EECE, MAST, MECH
Goals & Standards	ICM, IHG, ION, MSGS	ICMB, ICM, IHG, ION, MSGS ECLS CEAM, EECE
Professional Development	IHG, IHS SAPL, SGPS, SLAW CEAM, AFRD	ICMB, IHG, MSGS, IHS SGPS, SLAW CEAM, CIVG, EECE, AFRD
Roles & Responsibilities	ICM, ION SMLS, SLAW COMP, CHEM	ICM, ION SELS, SMLS, SLAW COMP, CHEM, EECE
Teaching Opportunities	ICMB, IAH, IHG SAPL, SGPS, SMLS, NUBS CHEM, CIVG, COMP, AFRD, MATH, NIRES	ICMB, IAH, IHG, MSGS, NICR SGPS, SMLS, NUBS CEAM, CHEM, CIVG, COMP, EECE, MATH

Table 4: Schools/Institutes whose (adjusted) mean scores were significantly higher or lower than (i) their (adjusted) Faculty average and (ii) the (adjusted) University average, as shown in Figures 11 and 12. Those Schools/Institutes not mentioned have (adjusted) mean scores that are *not* significantly different from their (adjusted) Faculty mean/University mean.

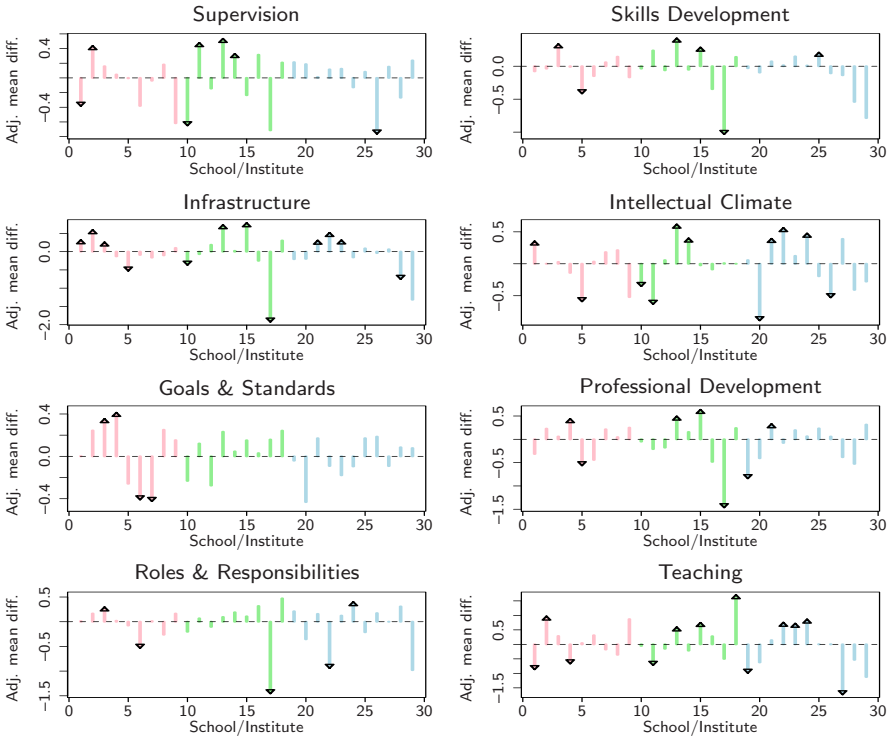

Figure 11: Adjusted mean differences from Faculty averages. △ indicates that the School/Institute (adjusted) mean is significantly *higher* than the (adjusted) Faculty mean; ▽ indicates that the School/Institute (adjusted) mean is significantly *lower* than the (adjusted) Faculty mean. Colours indicate Schools/Institutes belonging to **FMS**, **HaSS** and **SAgE**.

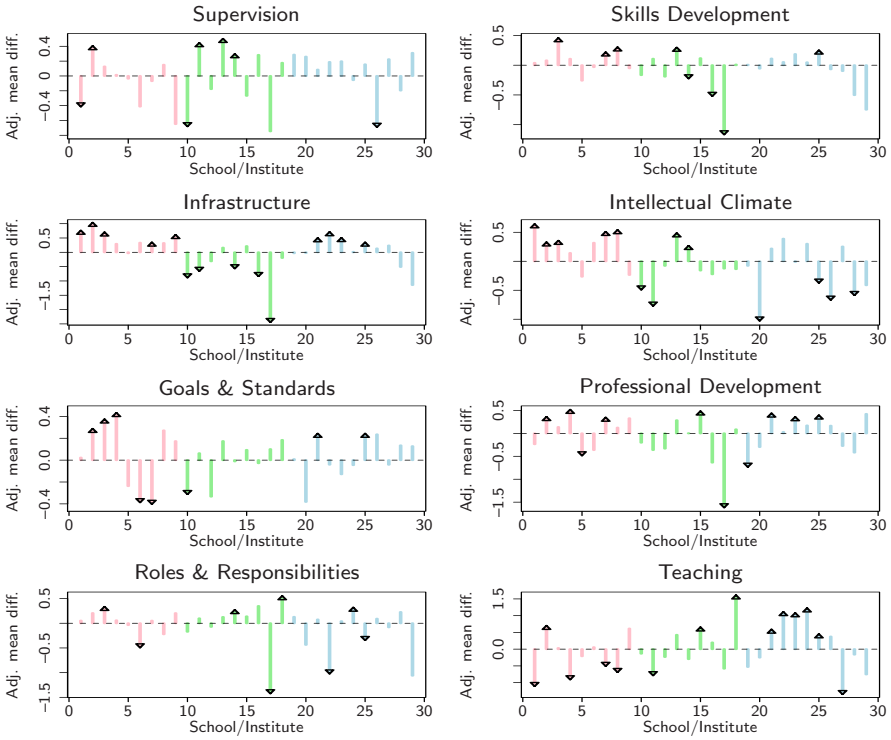

Figure 12: Adjusted mean differences from University averages. △ indicates that the (adjusted) School/Institute mean is significantly *higher* than the (adjusted) University mean; ▽ indicates that the (adjusted) School/Institute mean is significantly *lower* than the (adjusted) University mean. Colours indicate Schools/Institutes belonging to **FMS**, **HaSS** and **SAgE**.

5.3 Illuminating examples

Example 1
It turns out that PT students tend to give significantly lower scores for questions in section 6 of the PRES ("Professional Development & Careers") than do FT students: the average scores for PT and FT students here are 3.20 and 3.46 (respectively), and the corresponding p–value from a two sample t–test is 0.026. AFRD has a rather high proportion of its PT students not responding to the PRES (83% *c.f.* a School/Institute average of 42% of PT students not responding). Thus, we speculate that the mean score for "Professional Development & Careers" for AFRD should really be lower than that observed to account for the relatively high proportion of PT students not responding. Indeed, the adjusted mean for this section of the PRES is 2.28, compared to an unadjusted mean of 3.27 obtained directly from the raw data.

Example 2
Students over the age of 40 give, on average, significantly higher scores for section 5 of the PRES ("Goals & Standards") than do those students no older than 40: average scores for those over 40, and those no older than 40, are 4.18 and 3.98 (respectively), giving a p–value of 0.029 when tested using a two sample t–test. Not a single student aged over 40 in IHG responded to the PRES; thus, we speculate that the mean score for "Goals & Standards" for IHG should really be a bit higher than that observed. In fact, our adjusted mean for this section of the PRES increases to 4.27, from a raw mean of 4.22.

Key findings

The results in Table 4 and Figures 11 and 12 are analogous to those given earlier in Table 1 and Figures 3 and 4 (respectively), but now a statistical procedure has been used to adjust mean scores for each area of the key set, by School/Institute, to account for bias in non–response. For example, we find that the older a student, the less likely they are to complete the PRES; females are more likely to respond than males (significantly so if they are full time and younger than about 45); full time students are significantly more likely to complete the survey than part time students. The "Illuminating examples" given in section 5.3 serve to explain why it might be appropriate to adjust scores given this information. Thus, results in Table 4 should be seen as more reliable than those in Table 1 when comparing the performance of Schools/Institutes to their (adjusted) Faculty averages, and the (adjusted) University averages as a whole.

In terms of analysing trends over time (section 3) and comparing Newcastle's results this year with other institutions (section 4), we do *not* use scores that have been adjusted for bias. Ideally, we *would*. However, as explained in section 5.1.2, in order to implement this adjustment we need full demographic information on *all* students invited to complete each survey (whether they completed or not); for previous surveys at Newcastle (2008, 2009) and for this year's surveys at other institutions (e.g. other Russell Group institutions) we do not have such information.

Performing the adjustments for bias in non–response, as we do in this section, should be the way future postgraduate research experience surveys (and indeed other surveys!) at Newcastle (and other institutions!) are analysed, enabling a like–for–like comparison with the adjusted results from this year's survey reported in this section.

Acknowledgements

We thank Jonathon Gray, an administrator in Student and Academic Services, for providing data on those students who did not complete the PRES.

Lee Fawcett
Lecturer in Statistics
School of Mathematics & Statistics
Herschel Building
lee.fawcett@ncl.ac.uk

Appendix

	Univ.	FMS	IAH	ICMB	ICM	IHG	IHS	ION	MSGS	NICR	PSYC
Sample size	*919**	*312*	*32*	*39*	*50*	*26*	*32*	*23*	*74*	*26*	*10*
Supervision	4.23 4.50 (3.83, 4.83)	4.26 4.33 (4.00, 4.83)	4.18 4.17 (3.79, 4.58)	4.43 4.58 (4.04, 4.83)	4.33 4.33 (4.00, 4.83)	4.15 4.17 (3.83, 5.00)	4.10 4.08 (3.67, 4.92)	4.00 4.42 (3.38, 4.75)	4.34 4.50 (4.00, 4.83)	4.53 4.67 (4.33, 5.00)	3.63 4.00 (3.17, 4.50)
Skills Development	4.04 4.00 (3.67, 4.50)	4.10 4.00 (3.67, 4.67)	4.11 4.17 (3.88, 4.62)	4.23 4.00 (3.67, 4.67)	4.22 4.17 (3.83, 4.79)	4.18 4.08 (3.96, 4.67)	4.02 4.00 (3.54, 4.46)	3.78 3.83 (3.58, 4.08)	4.03 4.00 (3.67, 4.67)	4.35 4.50 (4.12, 4.83)	3.77 3.83 (3.54, 4.29)
Infrastructure	4.05 4.17 (3.67, 4.67)	4.27 4.33 (3.83, 4.83)	4.26 4.17 (4.00, 4.83)	4.48 4.50 (4.17, 5.00)	4.28 4.33 (3.83, 4.67)	4.18 4.00 (3.92, 4.75)	4.07 4.00 (3.67, 4.83)	4.13 4.33 (3.67, 4.83)	4.32 4.33 (4.00, 4.83)	4.20 4.33 (3.83, 5.00)	4.35 4.17 (4.00, 4.83)
Intellectual Climate	3.76 3.80 (3.20, 4.40)	3.96 4.00 (3.40, 4.60)	4.21 4.20 (3.90, 4.60)	3.80 3.80 (3.40, 4.40)	4.11 4.00 (3.80, 4.80)	3.85 4.00 (3.60, 4.20)	3.51 3.50 (3.05, 4.00)	4.06 4.00 (3.40, 4.80)	4.06 4.00 (3.40, 4.90)	4.02 4.20 (3.80, 4.60)	3.71 3.60 (3.40, 4.40)
Goals & Standards	3.99 4.00 (3.50, 4.75)	4.02 4.00 (3.50, 4.75)	3.95 4.00 (3.50, 4.56)	4.13 4.00 (3.75, 5.00)	4.34 4.50 (4.00, 5.00)	4.22 4.00 (4.00, 4.94)	3.83 3.75 (3.38, 4.50)	3.66 3.75 (3.31, 4.00)	3.80 3.75 (3.25, 4.50)	4.25 4.25 (4.00, 5.00)	4.03 4.12 (3.25, 4.69)
Professional Development	3.43 3.67 (3.00, 4.00)	3.46 3.67 (3.00, 4.00)	3.16 3.00 (3.00, 4.00)	3.50 3.33 (3.00, 4.00)	3.55 3.50 (3.00, 4.00)	3.62 3.83 (3.25, 4.00)	3.57 3.67 (3.00, 4.00)	2.83 3.00 (2.00, 4.00)	3.55 4.00 (3.00, 4.00)	3.42 4.00 (3.00, 4.00)	4.17 4.83 (4.08, 5.00)
Roles & Responsibilities	3.87 4.00 (3.50, 4.25)	3.98 4.00 (3.50, 4.50)	3.88 4.00 (3.25, 4.25)	4.08 4.00 (3.75, 4.50)	4.13 4.00 (3.69, 4.75)	4.08 4.00 (3.88, 4.50)	3.81 3.75 (3.31, 4.00)	3.59 3.75 (3.12, 4.00)	3.98 4.00 (3.75, 4.50)	3.98 3.88 (3.50, 4.62)	4.14 4.25 (3.50, 4.75)
Teaching	3.71 4.00 (3.00, 4.67)	3.77 4.00 (3.33, 4.67)	3.15 3.67 (2.83, 3.83)	4.15 4.33 (4.00, 5.00)	4.33 4.67 (4.00, 5.00)	3.23 3.33 (2.50, 4.00)	3.94 4.00 (3.33, 4.42)	3.58 3.67 (2.83, 4.17)	3.59 3.83 (3.00, 4.33)	3.11 3.67 (2.00, 4.33)	4.50 4.83 (4.42, 5.00)

Table 5: Faculty of Medical Sciences (FMS): PRES summaries by institute. In each cell are response mean scores (left), median scores (right) and interquartile ranges (in parentheses). *926 students responded, but 7 failed to give their Faculty or School/Institute.

	Univ.	HaSS	ECLS	NUBS	SACS	SAPL	SELS	SGPS	SHIS	SLAW	SMLS
Sample size	*919**	*322*	*70*	*38*	*31*	*31*	*79*	*34*	*19*	*10*	*10*
Supervision	4.23 4.50 (3.83, 4.83)	4.26 4.50 (3.83, 5.00)	4.07 4.17 (3.50, 4.83)	4.32 4.83 (3.75, 5.00)	3.93 4.25 (3.50, 4.83)	4.39 4.67 (4.08, 4.92)	4.40 4.67 (4.17, 5.00)	4.34 4.50 (4.08, 4.75)	4.33 4.50 (4.08, 5.00)	4.38 4.75 (4.50, 5.00)	4.10 4.17 (4.00, 4.46)
Skills Development	4.04 4.00 (3.67, 4.50)	3.97 4.00 (3.67, 4.50)	3.81 3.83 (3.33, 4.33)	4.09 4.17 (3.79, 4.67)	3.74 3.83 (3.67, 4.33)	4.18 4.17 (3.75, 4.83)	4.08 4.17 (3.67, 4.67)	4.09 4.00 (3.67, 4.50)	3.83 4.00 (3.67, 4.17)	3.88 3.75 (3.62, 4.71)	3.80 4.00 (3.46, 4.12)
Infrastructure	4.05 4.17 (3.67, 4.67)	3.76 3.83 (3.33, 4.50)	3.51 3.50 (3.00, 4.08)	4.02 4.00 (3.58, 4.62)	3.62 4.00 (2.83, 4.50)	4.11 4.17 (3.67, 4.50)	3.81 3.83 (3.50, 4.50)	3.98 4.00 (3.75, 4.50)	3.42 3.42 (2.79, 4.04)	3.20 3.17 (2.33, 3.33)	3.61 3.50 (3.33, 4.33)
Intellectual Climate	3.76 3.80 (3.20, 4.40)	3.65 3.80 (3.00, 4.40)	3.34 3.40 (2.60, 4.20)	3.54 3.70 (3.00, 4.25)	3.49 3.60 (3.05, 4.20)	3.88 4.00 (3.40, 4.60)	3.96 4.20 (3.60, 4.75)	3.59 3.60 (3.15, 4.25)	4.06 4.40 (3.40, 5.00)	3.30 3.40 (3.25, 3.75)	3.30 3.50 (2.70, 3.75)
Goals & Standards	3.99 4.00 (3.50, 4.75)	3.99 4.00 (3.50, 5.00)	3.76 4.00 (3.25, 5.00)	4.02 4.00 (3.50, 5.00)	3.65 3.88 (3.25, 4.19)	4.15 4.00 (4.00, 4.81)	4.08 4.00 (3.50, 5.00)	4.20 4.12 (3.75, 4.81)	4.14 4.12 (3.75, 4.94)	4.05 4.00 (4.00, 4.38)	4.05 4.00 (4.00, 4.44)
Professional Development	3.43 3.67 (3.00, 4.00)	3.42 3.67 (3.00, 4.00)	3.41 3.33 (2.83, 4.17)	3.45 3.67 (3.00, 4.00)	2.95 3.00 (2.67, 3.67)	3.47 3.67 (3.00, 4.00)	3.56 3.67 (3.00, 4.17)	3.50 3.67 (3.00, 4.33)	3.76 4.00 (2.75, 4.92)	3.00 3.00 (2.00, 4.00)	3.10 3.17 (2.75, 4.00)
Roles & Responsibilities	3.87 4.00 (3.50, 4.25)	3.85 4.00 (3.50, 4.50)	3.62 3.75 (3.19, 4.25)	3.97 4.00 (3.38, 4.62)	3.59 3.50 (3.25, 4.25)	4.01 4.00 (3.75, 4.25)	4.01 4.00 (3.50, 4.88)	3.73 3.75 (3.50, 4.00)	4.07 3.88 (3.75, 4.94)	3.61 3.75 (3.12, 4.12)	4.15 4.12 (3.81, 4.62)
Teaching	3.71 4.00 (3.00, 4.67)	3.68 4.00 (3.00, 4.67)	3.42 4.00 (2.67, 4.33)	3.20 3.33 (2.42, 4.00)	3.44 3.50 (3.00, 4.25)	3.46 4.00 (2.67, 4.25)	4.13 4.33 (3.92, 5.00)	4.12 4.33 (4.33, 4.67)	4.30 4.33 (4.00, 5.00)	3.00 3.17 (1.75, 4.42)	3.85 4.33 (3.67, 4.67)

Table 6: Faculty of Humanities and Social Sciences (HaSS): PRES summaries by School. In each cell are response mean scores (left), median scores (right) and interquartile ranges (in parentheses). *926 students responded, but 7 failed to give their Faculty or School/Institute.

	Univ.	SAgE	AFRD	BIOL	CEAM	CHEM	CIVG	COMP	EECE	MAST	MATH	MECH	NIRES
Sample size	*919**	*285*	*22*	*20*	*46*	*14*	*48*	*29*	*56*	*19*	*12*	*13*	*6*
Supervision	4.23 4.50 (3.83, 4.83)	4.20 4.33 (3.83, 4.83)	4.37 4.50 (4.17, 5.00)	4.41 4.67 (4.00, 5.00)	4.19 4.42 (3.83, 4.83)	4.40 4.50 (4.00, 5.00)	4.09 4.33 (3.79, 4.67)	4.26 4.33 (3.88, 5.00)	4.24 4.33 (3.88, 4.83)	4.13 4.42 (3.50, 4.50)	4.11 4.17 (3.75, 4.75)	3.76 4.00 (3.17, 4.67)	4.25 4.25 (4.00, 4.62)
Skills Development	4.04 4.00 (3.67, 4.50)	4.06 4.00 (3.67, 4.50)	4.12 3.92 (3.67, 4.67)	4.37 4.50 (4.00, 4.83)	3.97 4.00 (3.67, 4.33)	4.22 4.17 (4.00, 4.67)	4.04 4.17 (3.67, 4.50)	4.09 4.00 (3.83, 4.42)	4.15 4.08 (3.83, 4.50)	4.01 4.00 (3.33, 4.50)	3.69 3.75 (3.42, 4.17)	3.89 4.17 (3.96, 4.33)	3.77 3.50 (3.50, 3.67)
Infrastructure	4.05 4.17 (3.67, 4.67)	4.10 4.17 (3.83, 4.67)	3.87 4.00 (3.58, 4.50)	4.14 4.00 (3.67, 4.83)	4.15 4.17 (3.83, 4.67)	4.46 4.67 (4.00, 4.83)	4.19 4.33 (3.83, 4.83)	4.18 4.17 (3.83, 4.67)	4.15 4.08 (3.83, 4.62)	3.71 3.75 (3.04, 4.29)	4.33 4.50 (4.04, 4.67)	3.67 3.83 (3.67, 3.83)	3.58 3.83 (3.25, 4.17)
Intellectual Climate	3.76 3.80 (3.20, 4.40)	3.71 3.80 (3.20, 4.20)	3.63 3.80 (3.15, 4.05)	3.29 3.30 (2.70, 4.00)	3.84 4.00 (3.55, 4.40)	3.81 4.00 (3.20, 4.15)	3.64 3.80 (3.00, 4.20)	4.10 4.10 (3.50, 4.95)	3.75 3.60 (3.20, 4.20)	3.52 3.60 (3.00, 4.30)	3.70 3.90 (3.20, 4.25)	3.45 3.60 (3.05, 4.05)	3.56 3.80 (3.20, 4.00)
Goals & Standards	3.99 4.00 (3.50, 4.75)	4.00 4.00 (3.50, 4.75)	3.95 4.00 (3.75, 4.25)	4.12 4.00 (4.00, 4.75)	4.05 4.00 (3.75, 4.50)	3.82 4.00 (3.56, 4.00)	3.94 4.00 (3.50, 4.50)	4.04 4.00 (3.50, 5.00)	4.05 4.00 (3.94, 4.75)	4.01 4.00 (3.25, 4.75)	3.92 4.00 (3.44, 4.81)	3.87 4.00 (3.50, 4.75)	3.96 3.88 (3.75, 4.00)
Professional Development	3.43 3.67 (3.00, 4.00)	3.42 3.33 (3.00, 4.00)	3.27 3.17 (2.92, 4.00)	3.67 3.50 (3.00, 4.08)	3.55 3.67 (3.00, 4.00)	3.18 3.33 (3.00, 3.33)	3.33 3.33 (3.00, 4.00)	3.54 3.50 (3.00, 4.00)	3.50 3.67 (3.00, 4.00)	3.58 4.00 (3.00, 4.00)	2.69 2.50 (1.50, 4.00)	3.33 3.33 (2.83, 4.00)	3.33 3.50 (2.58, 3.92)
Roles & Responsibilities	3.87 4.00 (3.50, 4.25)	3.82 4.00 (3.25, 4.25)	3.80 3.75 (3.50, 4.00)	3.86 4.00 (3.44, 4.12)	3.85 4.00 (3.50, 4.12)	3.60 3.75 (3.31, 4.00)	3.75 4.00 (3.25, 4.25)	4.07 4.00 (3.50, 4.62)	3.84 4.00 (3.25, 4.25)	3.72 3.88 (3.06, 4.25)	3.73 3.75 (3.38, 4.00)	3.73 3.75 (3.25, 4.50)	3.81 3.62 (3.38, 4.06)
Teaching	3.71 4.00 (3.00, 4.67)	3.68 4.00 (3.00, 4.33)	3.11 3.50 (2.25, 4.00)	3.70 4.00 (3.17, 4.50)	3.34 3.33 (2.67, 4.17)	3.53 3.67 (3.00, 4.08)	3.98 4.00 (3.67, 4.67)	3.97 4.33 (3.33, 4.67)	3.78 4.00 (3.00, 4.33)	3.76 4.00 (3.00, 4.50)	3.22 3.00 (3.00, 3.33)	3.90 4.00 (3.83, 4.33)	2.78 3.00 (2.67, 3.00)

Table 7: Faculty of Science, Agriculture and Engineering (SAgE): PRES summaries by School. In each cell are response mean scores (left), median scores (right) and interquartile ranges (in parentheses). *926 students responded, but 7 failed to give their Faculty or School/Institute.