Image interpretation performance : a longitudinal study from novice to professional

WRIGHT, Chris and REEVES, Pauline

Published version

WRIGHT, Chris and REEVES, Pauline (2016). Image interpretation performance : a longitudinal study from novice to professional. Radiography.

Copyright and re-use policy

Title:

Image Interpretation Performance: A Longitudinal Study from Novice to Professional

Corresponding Author:

Dr Chris Wright

Sheffield Hallam University

Sheffield

England

S10 2BP

chris.wright@shu.ac.uk

+44 144 225 5488

Co-Author:

Dr Pauline Reeves (address as above)

Keywords:

RadBench, Image Interpretation, Red Dot, Radiographer, PCE, Clinical governance

Wordcount:

2303

Highlights

1. Some novices appear to have inherent skills in fracture identification.
2. RadBench testing as part of the UCAS selection process provides a useful indication of future performance.
3. Increase in specificity is the primary gain of university education in increasing accuracy.
4. Radiography graduates may require further education in order to deliver reliable image interpretation decisions.

Abstract 248

Purpose

Universities need to deliver educational programmes that create radiography graduates who are ready and able to participate in abnormality detection schemes, ultimately delivering safe and reliable performance because junior doctors are exposed to the risk of misdiagnosis if unsupported by other healthcare professionals. Radiographers are ideally suited to this role having the responsibility for conducting the actual X-ray examination.

Method

The image interpretation performance of one cohort of student radiographers was measured upon enrolment from UCAS in the first week of university education and then again prior to graduation using RadBench (n=23).

Results

The results identified that novices have a range of natural image interpretation skills; accuracy 35-85%, sensitivity 45-100%, specificity 15-85%, mean ROC 0.691. Graduates presented a narrower range; accuracy 60-90%, sensitivity 40-100%, specificity 60-90%, mean ROC 0.841. The positive shift in graduate mean accuracy (+16%) was driven by increases in specificity (+27%) rather than sensitivity (+5%). No statistically significant differences (ANOVA) could be found between age group, gender and previous education however trends were identified. One graduate could meet a 90% benchmark accuracy standard and 52% of the population (n=12) an 80% standard.

Conclusion

Image interpretation testing at the point of UCAS entry is a useful indicator of future performance and is a recommended factor for consideration as part of the selection process. Whilst image interpretation now forms an integral part of undergraduate radiography programmes, new graduates may not necessary possess the reliability in decision making to justify participation in abnormality detection schemes, highlighting the need for continuous professional development.

Introduction 589

This paper presents a longitudinal study of the image interpretation skills of student radiographers from enrolment to graduation and considers the implications for the profession and the NHS in terms of reliable abnormality detection to aid service improvement in Accident and Emergency (A&E) departments.

Traditional UK National Health Service (NHS) practice is for a patient presenting in A&E be seen by a doctor, referred for X-ray, and then return to the doctor for evaluation. Rotation through A&E presents an important potential learning and development opportunity for junior doctors. Their lack of radiological expertise and related knowledge, however, exposes them to the risk of misdiagnosis if unsupported by other healthcare professionals. One solution might be to increase the number of radiologists in order to provide immediate reporting of images. This presents two key challenges, firstly, the demand for diagnostic imaging services has grown faster than the supply and, secondly, the high fiscal cost. The potential for radiographers to deliver equivalent accuracy of reporting to radiologists[1] offers an alternative solution. The joint publication of the Royal College of Radiologists and the College of Radiographers[2] takes a team working approach to formal image reporting, recognising the value of radiographers in delivering timely decisions ('hot reporting') to support patient management. The 2008 scope of practice survey[3] identified that 53% of participating NHS sites employed reporting radiographers. Hot reporting is generally only available during the day but a few centres offer it at night.

Initial image interpretation may also be performed by the examining radiographer, with a formal report provided either by a radiologist or reporting radiographer at a later stage. Radiography abnormality detection schemes (RADS) have traditionally focussed around 'red dot' and have been used for over 25 years[4]. There has recently been a push towards to the use of preliminary clinical evaluation (PCE), also known as 'commenting'. A UK wide survey[4] identified that 93% of participating hospitals operated abnormality signalling systems, although only 25% considered this to be a mandatory function.

A proliferation of studies[5,6,7,8,9] have highlighted deficiencies in the image interpretation competence of medical students and attributed this to the lack of formal radiological tuition. Unlike medical degree programmes, which offer limited exposure to formal instruction in X-ray image interpretation as part of undergraduate education[10], modern undergraduate diagnostic radiography degree programs have changed[11] to meet the College of Radiographer's[12] policy that expects graduates to be able to provide reliable preliminary clinical evaluation (PCE) based on the radiographic images that they produce. The aim of this is to provide the referring doctor with key information to underpin the diagnostic decision. Whilst 'red dot' has enabled radiographers to make contributions to A&E services for many years, the College of Radiographers[12] argued that this approach no longer aligned with current clinical governance processes and should be phased out and replaced by PCE such that junior doctors would be provided with more directed information on which to base their patient treatment decisions. The first step in scaffolding this transition is developing the ability of radiographers to make the correct image interpretation decision before increasing confidence and then learning to write the PCE[13]. With further development some of these radiographers would then form the pool of future reporters.

Novices enrolling on diagnostic degree programmes in principle all start from the same point and undergo the same opportunities for image interpretation education within the same university; however in practice this might not actually be the case. This research study aimed to measure the performance of one full cohort of radiography students from a single university at the point of enrolment onto the undergraduate course from UCAS and compare it to their exit performance upon graduation.

Method 254

The decision-making performance of a single cohort of student radiographers at one university was measured at the point of enrolment from UCAS in week one of the first year and again one month prior to graduation in year three using the abnormality signalling component of RadBench, a specifically developed software program for measuring image interpretation peformance[14]. Cognisant that making the correct decision is a precursor to accurate written description[13], the option to collect preliminary clinical evaluation (PCE) was disabled in order to focus directly on decision making. The research received ethical approval from the study university. Students were provided with a participant information sheet and gave their written consent (n=36). Participation was voluntary.

The test bank contained twenty appendicular musculo-skeletal images (see figure 1) which had a fifty per cent incidence of abnormality, confirmed by prior blind double reporting. Images were selected such that abnormality was restricted to a single fracture per image, all clearly visible with satisfactory search. Respondents were asked to choose from five options per image that best described their decision making confidence (1=Definitely Normal, 2= Probably Normal, 3=Possibly Abnormal, 4= Probably Abnormal, 5= Definitely Abnormal). This format enables identification of decision making confidence and also facilitates the calculation of ROC. For the calculation of sensitivity, specificity and accuracy this data is binarised into normal and abnormal decisions. The distribution of images is illustrated in figure 1.

<insert figure 1 - Case Mix>

An identical randomised image bank was used for both tests. Answers were not revealed after the enrolment test. Students were unaware that the graduation test was a randomised clone of the enrolment test.

The results from both tests were analysed in terms of accuracy, sensitivity, specificity to compare enrolment with graduation performance; analysis of variance (ANOVA) with previous education, gender and age group. The receiver operator characteristic (ROC) was calculated with the JROCFIT web based calculator[15].

Results 405

Thirteen students elected not to complete the final assessment and were therefore excluded from the data analysis (n=23). Their demographics are presented in figure 2.

<insert figure 2 – Population Demographics>

Figure 3 provides a box-plot to summarise performance.

<insert figure 3>

Mean sensitivity at enrolment was 73% (std dev=0.157) with a range from 45 to 100%. Mean sensitivity at graduation was 78% (std dev=0.107) with a range from 40 to 100%.

Mean specificity at enrolment was 49% (std dev=0.153) with a range from 15 to 85%. Mean specificity at graduation was 76% (std dev=0.153) with a range from 40 to 100%.

Mean accuracy at enrolment was 61% (std dev=0.113) with a range from 35 to 85%. Mean accuracy at graduation was 77% (std dev=0.072) with a range from 60 to 90%.

ROC at enrolment was 0.691increasing to 0.841 at graduation. See figure 4

<insert figure 4>

Figure 5 demonstrates the difference in accuracy between enrolment from UCAS and graduation by student. Supporting the evidence of the ROC, the accuracy of 91% (n=21) of students improved, one stayed the same, and one decreased. Analysis of variance (ANOVA) demonstrated no significant differences at a 95% confidence level (p=0.555).

<insert figure 5>

The mean accuracy improvement was 16%, driven predominantly by the 27% increase in specificity relative to the sensitivity which increased by only 5%.

Unsurprisingly no student could meet a 90% benchmark standard upon enrolment from UCAS although 13% (n=3) could achieve the 80% standard. At graduation 4% (n=1) could meet a 90% standard and 52% (n=12) an 80% standard.

Considering graduate performance, analysis of variance (ANOVA) demonstrated no significant differences at a 95% confidence level between gender (p=0.370), age group (p=0.919) or previous education (p=0.137) although the box-plots do indicate a trend (see figures 5-7). Males tended to deliver higher performance accuracy within this population. The median performance of all groups, except 31-45, was almost identical. The 31-45 age group population did however present a narrower performance distribution range. The range in accuracy performance is wide. The overall trend is seemingly that A-level students and the one with a previous degree tended to perform to a higher level at the point of graduation, relative to BTEC and Access entry students.

Figure 6 considers the accuracy of graduates by gender.

<insert figure 6>

Figure 7 considers the accuracy of graduates relative to age group.

<insert figure 7>

Figure 8 considers the accuracy of graduates relative to their education prior to university entry.

<insert figure 8>

Discussion 891

This research has explored the changing image interpretation capabilities of novices and radiography graduates. It is perhaps noteworthy that the mean accuracy (61%) at the point of entry to higher education was far higher than anticipated, with a range extending from 35-85%, suggesting that image perception comes perhaps more naturally to some than others; certainly the ability to search and correctly identify abnormality (sensitivity). Whilst the reliability of this longitudinal study is limited, being carried out with only one population of student radiographers in one university, the evidence suggests that gender, age group and educational profile all potentially impact on graduate image interpretation performance and the radiography profession of the future. Novices enrolling on diagnostic degree programmes start from the same point on the course, but they have very different profiles. Whilst the same academic learning opportunities are available to all students, the clinical experience will vary depending on the placement sites, although no statistically significant differences in image interpretation performance could be identified.

Students undertook academic image interpretation modules in all three years of their education. The accuracy of all bar two students improved between enrolment from UCAS and graduation. One stayed the same and one decreased. Both these students suffered adverse personal circumstances during the course which affected performance in all aspects of their work.

With the exception of one student, males tended to deliver higher performance accuracy within this population although gender bias is noted (6 males versus 17 females). This finding supports the notion that males on average have one standard deviation higher spatial intelligence quotient than females[16] however other research has found no difference by gender[17] and identified that visuo-spatial aptitude can be enhanced through teaching[18]. It is perhaps the spatial aptitude that is important and not the gender; a subject for further research. Another factor could be that the five top performing males formed part of the A-level entry group who also outperformed the other education modes of entry, although not to a statistically significant level. Whilst all A-level students studied biology as a mandatory requirement of the admissions process, the value of this logic is unclear when many courses had little 'human' content. The reasons why BTEC and Access entry students performed relatively less well are another subject of further research. The median performance of all groups, except those aged 31-45, was almost identical. This latter group presented a narrower performance distribution range with no very poor performers. A common route to higher education for mature students is through Access[19] as it provides a wide range of module options, often on a part-time basis, to build the desired educational portfolio. Students following the BTEC route tended to be of a similar age to the A-level entrants however the mode of assessment is very different. This 'no examinations' route is better suited to some students and enables them to achieve equitable UCAS points to the A-level route[20]. The image interpretation tests utilised in this research are, however, a form of examination and so this may have a negative impact on the results of the BTEC entrants.

Assuming entry qualifications and expected humanistic values are evident in applicants for diagnostic radiography programmes, image interpretation testing at the UCAS entry point might assist in selection. RadBench[21] has offered this option since 2011 for applicants to diagnostic radiography and medicine and many applicants offer their performance as evidence in their personal statements. The data supports the concept that, upon entry, candidates should have minimum accuracy of 50%[22] with a higher sensitivity than specificity because the ability to correctly identify normality is the key learning improvement over the duration of the university programme (27% mean gain) versus only 5% for sensitivity.

Neither the College of Radiographers nor the Health and Care Professions Council (HCPC) require a defined quality standard of image interpretation performance in order to maintain registration and assure safe practice. A minimum performance benchmark accuracy of 80% has been suggested[23] as this historically might be typical of radiographers and junior doctors, although educational practices have changed and it is perhaps now reasonable to expect radiographers to deliver 90% accuracy (one error in ten) in-line with the Fellowship of the Royal College of Radiologists (FRCR) Part B rapid reporting test[24] before participating in any form of abnormality signalling system. The actual measurement in practice is unknown for radiographers, doctors (or indeed any other healthcare professional) however RadBench offers the potential for benchmarking in the future. This research suggests that, whilst the university has now embedded image interpretation skills within the undergraduate degree programme in line with other institutions[11], only 56.5% of the population could meet the 80% benchmark including only one student the 90% benchmark at the point of graduation. It is possible however that these results contain the ‘tail off component‘[22] which is reported to consistently occur in the second semester of year three, and students could well have preformed to a higher standard earlier in their education. The reasons for the decline in performance are unclear and the subject of further research.

The requirement of the College of Radiographers[12] that 'red dot' signalling systems be replaced by written PCE should be considered in the context of need, quality and implementation. From the perspective of clinical governance, simply replacing red dot with PCE offers no quality improvement if the decisions that are made are unreliable. From an organisational perspective, in the management of A&E referrals, hot reporting probably offers the most robust and accurate decision making process, however requires additional funding to support this approach. Arguably it may also be unnecessary if the examining radiographers are able to provide high quality reliable abnormality signalling. Abnormality signalling by radiographers with later image reporting however presents two critical to quality issues; the first is the benchmark entry point and the second is the continuous monitoring of performance. The evidence from this research suggests that many graduates may be unable to achieve a minimum 90% accuracy and further continuous professional development will be required to further develop their image interpretation skills to the benchmark level.

Conclusion 164

This research study aimed to measure the performance of one full cohort of radiography students from a single university at the point of enrolment onto the undergraduate course from UCAS and compare it to their exit performance upon graduation. It has shown that whilst mean accuracy increases from 61% to 77% this improvement is not statistically significant. The main performance increase over the duration of the university programme is in specificity (27% mean gain versus only 5% for sensitivity) demonstrating how radiography education has developed the skills to correctly identify normality and differentiate normal variants from potential fractures.

Universities need to deliver educational programmes that create radiography graduates who are ready and able to participate in abnormality detection schemes, ultimately delivering safe and reliable performance because junior doctors are exposed to the risk of misdiagnosis if unsupported by other healthcare professionals. Radiographers are ideally suited to this role having the responsibility for conducting the actual X-ray examination.

This research has demonstrated that novices entering higher education through UCAS have a wide range of natural ability to identify fractures, some are extremely good, and that gender, age group and previous education potentially impact on graduate image interpretation performance. Testing at the point of entry is a useful indicator of future image interpretation performance and is a recommended factor for consideration as part of the selection process.

Whilst image interpretation now forms an integral part of undergraduate radiography programmes, new graduates may not necessarily possess the reliability in decision making to justify participation in abnormality detection schemes, highlighting the need for continuous professional development. Introducing preceptorship during the first year of qualification would enable graduates to further develop their image interpretation skills before participating in abnormality detection schemes.

References

1 Manning DJ. Evaluation of diagnostic performance in radiography. Radiography. 1998 Feb 28;4(1):49-60.
2 The Royal College of Radiologists and the Society and College of Radiographers (2012) Team working in clinical imaging. London.
3 University of Hertfordshire and The Institute for Employment Studies (2008) Scope of Radiographic Practice: A report compiled by the University of Hertfordshire in collaboration with the Institute for Employment Studies for the Society and College of Radiographers.
4 Snaith B, Hardy M. Radiographer abnormality detection schemes in the trauma environment—An assessment of current practice. Radiography. 2008 Nov 30;14(4):277-81.
5 Scheiner JD, Noto RB, McCarten KM. Importance of radiology clerkships in teaching medical students life-threatening abnormalities on conventional chest radiographs. Academic radiology. 2002 Feb 28;9(2):217-20.
6 Jeffrey DR, Goddard PR, Callaway MP, Greenwood R. Chest radiograph interpretation by medical students. Clinical radiology. 2003 Jun 30;58(6):478-81.
7 Dawes TJ, Vowler SL, Allen CM, Dixon AK. Education improves medical student performance in image interpretation. The British journal of radiology. 2014 Feb 13.
8 Eisen LA, Hegde A, Berger JS, Narasimhan M, Schneider RF. Competency in chest radiography: A comparison of Medical students, Residents and Fellows. CHEST Journal. 2005 Oct 1;128(4_MeetingAbstracts):343S-a.
9 Boutis K, Pecaric M, Seeto B, Pusic M. Using signal detection theory to model changes in serial learning of radiological image interpretation. Advances in health sciences education. 2010 Dec 1;15(5):647-58.
10 Belfield J. Using Gagnes theory to teach chest X ‐ ray interpretation. The clinical teacher. 2010 Mar 1;7(1):5-8.
11 Hardy M, Snaith B. Radiographer interpretation of trauma radiographs: Issues for radiography education providers. Radiography. 2009 May 31;15(2):101-5.

12 The College of Radiographers (2013) Preliminary Clinical Evaluation and Clinical Reporting by Radiographers: Policy and Practice Guidance 11 February
13 Wright,C. (2012) 'RadBench: Benchmarking Image Interpretation Performance'. UKRC. June. Liverpool: http://shura.shu.ac.uk/id/eprint/8432
14 Wright C, Reeves P, RadBench: Benchmarking image interpretation skills, Radiography (2016), http:// dx.doi.org/10.1016/j.radi.2015.12.010
15 Eng J. ROC analysis: web-based calculator for ROC curves. Baltimore: Johns Hopkins University [updated 2014 March 19; cited December 2015]. Available from: http://www.jrocfit.org.
16 Vandenberg,S. & Kuse,A. (1978). Mental rotations, a group test of three dimensional spatial visualisation. Perception & Motor Skills. 47. p.599-604
17 Robert M, Chevrier E. Does men's advantage in mental rotation persist when real three-dimensional objects are either felt or seen?. Memory & cognition. 2003 Oct 1;31(7):1136-45.
18 Lord TR. Enhancing the visuo-spatial aptitude of students. Journal of Research in Science Teaching. 1985 May 1;22(5):395-405.
19 Beetham H, Sharpe R. Rethinking pedagogy for a digital age: Designing for 21st century learning. Routledge; 2013 Apr 17. p85-86
20 https://www.ucas.com/ucas/undergraduate/getting-started/entry-requirements/tariff/calculator - accessed December 2015
21 RadBench (2014) http://www.radbench.org. Available via Papaya UK Ltd
22 Wright,C.(2014) 'RadBench : Benchmarking Image Interpretation Skills on a Global Scale. Achieving Excellence in Radiography Education & Research Conference. Bristol. Available at http://shura.shu.ac.uk/id/eprint/10240
23 The Royal College of Radiologists. 2016. Available at https://www.rcr.ac.uk/clinical-radiology/examinations/final-frcr-part-b-examination-0
24 Brealey S. Quality assurance in radiographic reporting: a proposed framework. Radiography. 2001 Nov 30;7(4):263-70.

FIGURES

Figure 1 – Case Mix

	Cases	Normal	Fracture
Ankle	4	2	2
Foot	4	2	2
Knee	2	1	1
Hand	4	2	2
Wrist	2	1	1
Elbow	2	1	1
Shoulder	2	1	1
Total	20	10	10

Figure 2 – Population Demographics

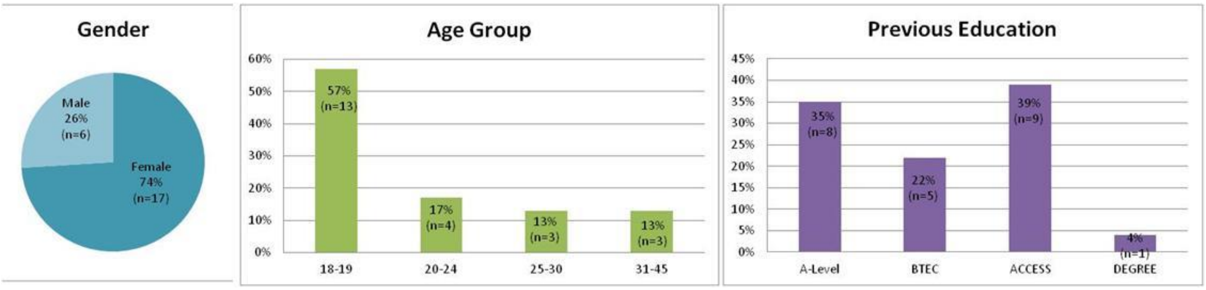

Figure 3 – UCAS Entry versus Graduate Performance

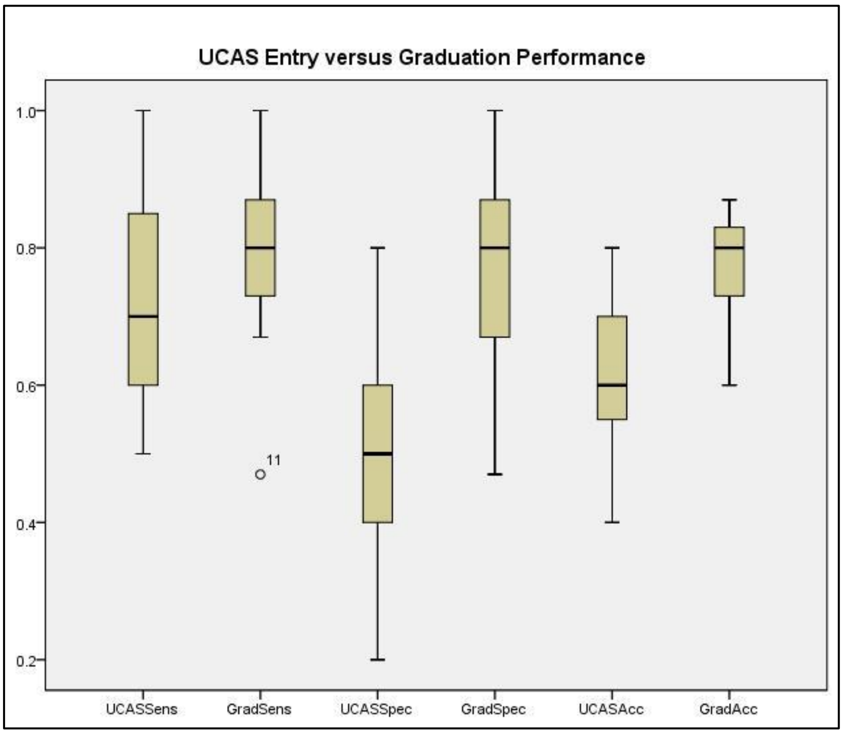

Figure 4 – ROC Analysis

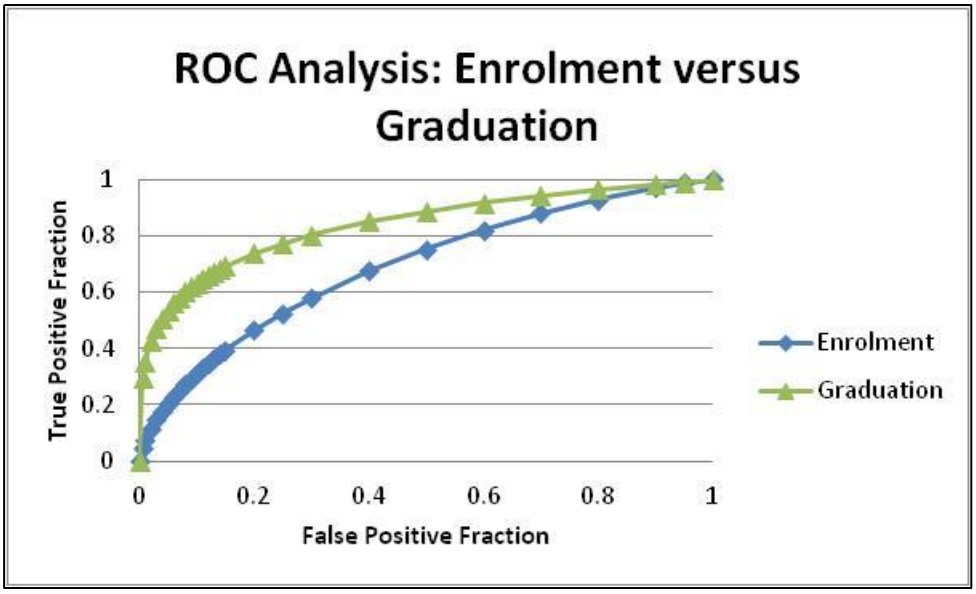

Figure 5 – UCAS versus Graduate Accuracy

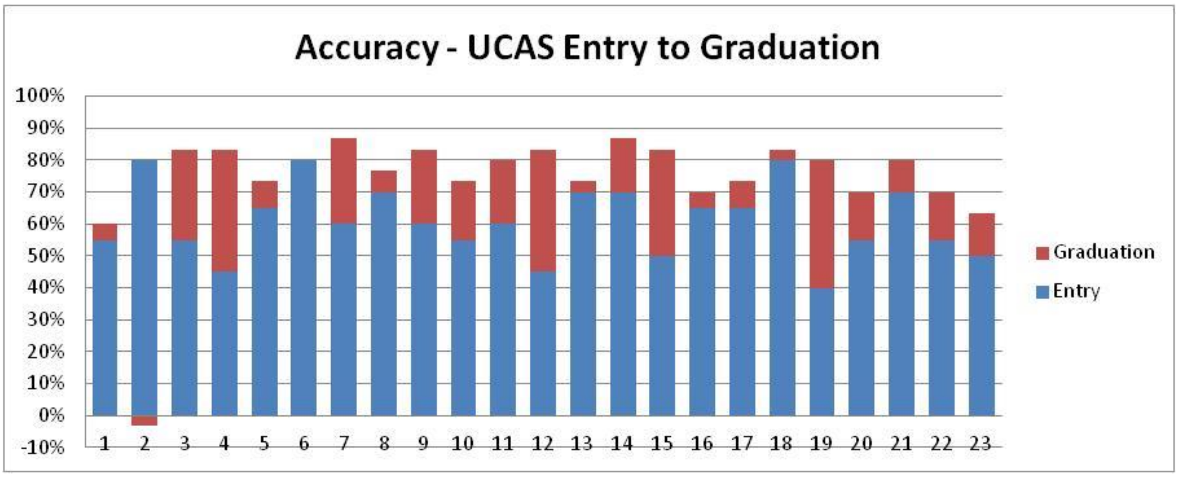

Figure 6 – Graduation Accuracy by Gender

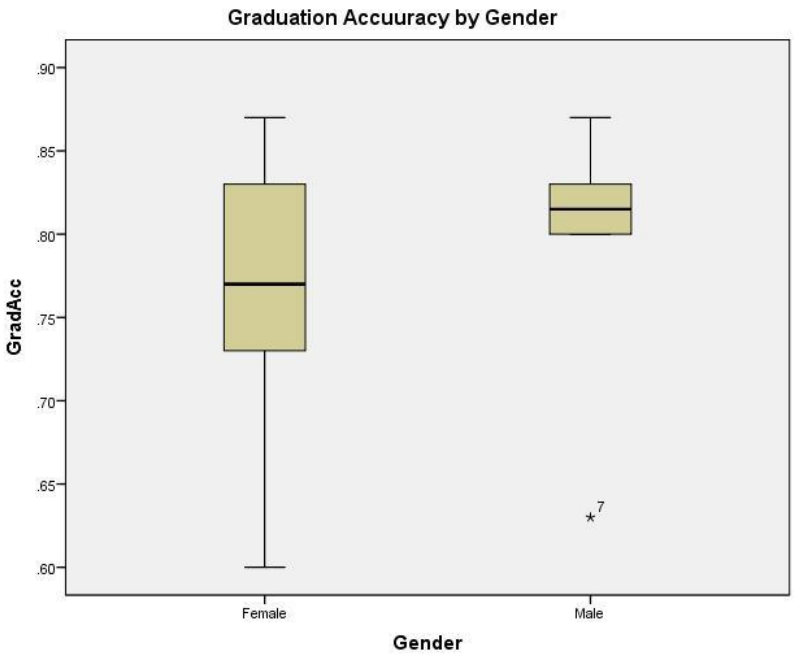

Figure 7 – Graduation Accuracy by Age Group

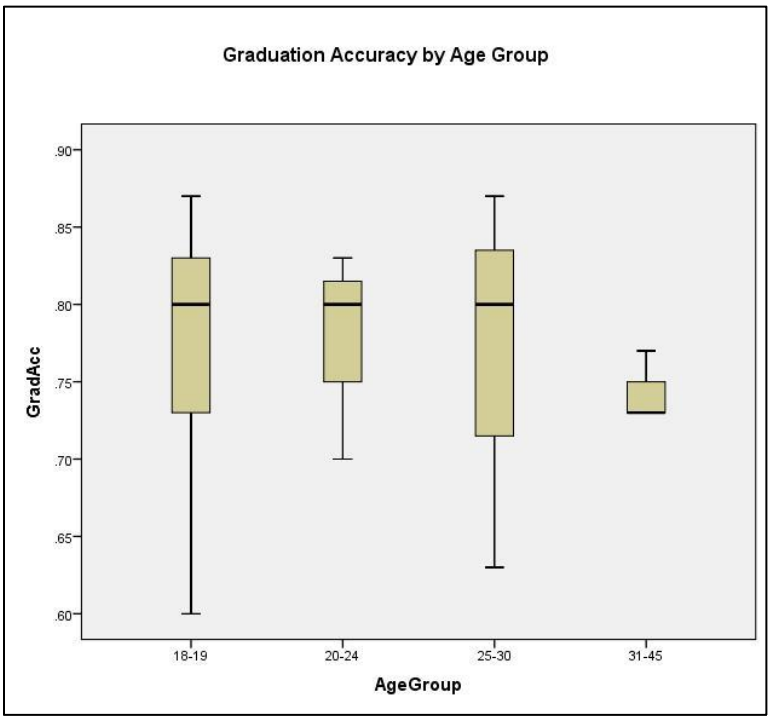

Figure 8 – Graduation Accuracy by Previous Education