CASE STUDY

Creating statistics e-assessments using Dewis with embedded R code

Iain Weir, Department of Engineering Design and Mathematics, University of the West of England, Bristol, U.K. Email: iain.weir@uwe.ac.uk
Rhys Gwynllyw, Department of Engineering Design and Mathematics, University of the West of England, Bristol, U.K. Email: rhys.gwynllyw@uwe.ac.uk
Karen Henderson, Department of Engineering Design and Mathematics, University of the West of England, Bristol, U.K. Email: karen.henderson@uwe.ac.uk

Abstract

We report on the creation of statistics e-assessments using the Dewis system with embedded R code. Dewis is a fully algorithmic open-source e-assessment system designed and developed at the University of the West of England, Bristol (UWE). Dewis' ability to communicate with the R programming language greatly facilitates the task of generating bespoke data and its subsequent analysis. This approach has allowed us to successfully test students' ability to perform involved statistical analyses on individual data sets and led to the creation of a suite of open access online e-learning modules on the UK national **stats**tutor website. Development of a Dewis-R interface allows the creation of sophisticated e-assessments solely by writing an R script file. The goal is to create a community of Dewis-R practitioners who will be able to author and share relevant, authentic and engaging statistics e-assessments that enrich the learning experience of students.

Keywords: statistics, SPSS, R, e-assessment, Dewis.

1. Background

Great strides have been made in the use of computer aided assessment for mathematics over recent years and many mathematics departments regularly use formative and summative e-assessments for their students (Sangwin, 2013). Such e-assessments can supply challenging practice in mathematical skills that, given a suitable platform and connection, students can access anytime or anywhere and obtain instantaneous and timely adaptive feedback. However, there has been far less progress on using e-assessment to test the ability to perform statistical analyses despite the fact that statistics is a component in many university courses.

Dewis (2012) is a fully algorithmic open-source e-assessment system designed and developed at the University of the West of England, Bristol (UWE). Originally designed to support and assess the learning of mathematics, it is currently used in the fields of business, computer science, nursing, engineering and mathematics. Implemented for the first time in 2007 the system is now well-established and in 2015/16 within UWE and partner institutions, Dewis was used for formative and summative tests to support over 3,500 students involving more than 50,000 assessment attempts.

A recent innovation by Weir, Gwynllyw and Henderson (2015) was to successfully embed R statistics code within Dewis to create e-assessments that could take advantage of the statistical functions available within R (2014). This approach led to the creation of complex statistics e-assessments that replicate full statistical analyses and report writing. These e-assessments have been used on level 2 (second year undergraduate) research skills modules, delivered to 850+ Business School students at UWE, since 2014. Students receive their own unique data set, generated by Dewis, which they are able to download directly into Excel, perform the necessary operation in SPSS before re-logging back into Dewis to submit their answers (Gwynllyw, Weir and Henderson, 2016).

The success of this approach led to the creation of an open access statistics resource which is freely available from the **stats**tutor site (*Dewis on **stats**tutor*, 2015). This resource has been made available under a Creative Commons licence by the authors of this article and reviewed by Dr Nadarajah Ramesh, University of Greenwich following a **sigma** Network Resource Development Grant (2015). The resource comprises a suite of e-assessments together with supporting materials that relate to the statistical activities involved in choosing and carrying out an appropriate one sample test for location (mean or median) on a randomly generated data set. Five e-assessment modules are available, as shown in figure 1, and these may be accessed independently or can be taken sequentially mimicking the flow of a full statistical analysis using the SPSS software package. On accessing the resource a new statistical data set may be generated or an existing data set used. Each module requires the data to have been downloaded to the SPSS statistical package, relevant analysis output obtained and a few questions answered to demonstrate understanding of the results. On submission, the e-assessment system marks the responses immediately and provides full bespoke feedback for inappropriate test choices as well as other incorrect analysis. Videos and instruction pamphlets are accessible as links from each e-assessment, which give clear instructions as to how to carry out the analyses and interpret results using SPSS. These additional resources, together with repeated use of the e-assessment modules, facilitates learning how to identify and employ the correct test on a variety of data sets.

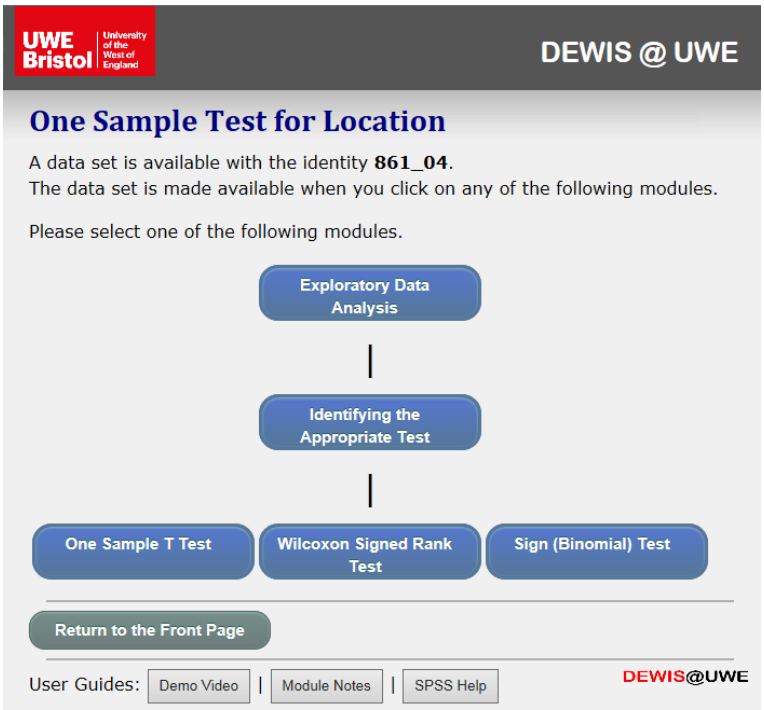


Figure 1. The five e-assessment modules that make up the resource on **stats**tutor.

2. Dewis-R Interface Project

The projects described in Section 1 demonstrate the potential of creating complex, intelligent statistics e-assessments using the Dewis system with embedded R code. However the drawback to expansion and growth of this approach has been in part the complexity of authoring questions. Not only do question authors have to be conversant with the R language they also have to be familiar with the programming of questions within Dewis. The issue of requiring familiarity with the Dewis syntax has been addressed by introducing a simplified Dewis question structure specific to this project. This simplification is possible since the statistical computations in these questions will be performed by R and not by Dewis. As such, Dewis supplies the administration of the assessment

but not the computations. This approach should appeal to statisticians who are already familiar with R. The question structure is in the form of a text file partitioned into two parts, the R part and the Dewis part. The Dewis part contains the syntax of the question as presented to the student, together with a declaration of the question parameters that will be used to present the question as well as to perform the marking and feedback. This part of the question is interpreted on installation of the question. On installation of the question, the R part of the question is written as an R script into the Dewis server. This script will be executed on every occasion the assessment is run. An execution of the R script will calculate the question parameters (including the answers) and pipe these to Dewis, enabling Dewis to present the question to the student. Weir and Gwynllyw have been awarded a UWE Teaching and Learning grant to push this idea forward and to create a cross-faculty statistics e-assessment Learning & Teaching community. The goal is to create a community of Dewis-R practitioners who will be able to author and share relevant, authentic and engaging statistics e-assessments that enrich the learning experience of students.

In this case study we demonstrate the Dewis-R interface by showing two examples. It is hoped that this will generate interest from the wider statistics community to trial this new approach of authoring. The examples presented here and others can be found at the Dewis-R Statistics Resources website (2016).

2.1. Example 1: Calculating and comparing a mean

The first example is by design very simple enabling us to demonstrate a basic Dewis-R script. The example involves the calculation of a mean from a sample of IQ scores and its subsequent comparison of it with the theoretical population mean of 100. The student is required to download an Excel file that contains a random data set to analyse. The sample mean is required to be given to one decimal place and from a dropdown menu the student has to report whether it is higher than, lower than or the same as the population average of 100. Thus this e-assessment comprises solely of one numerical input and one dropdown input. Figure 2 shows a screenshot of the question that is presented and, for illustration purposes, is of a student's attempt that has one correct and one incorrect answer. Note that in the subsequent feedback we colour code correct answers green and incorrect answers red, and hence have used this convention here.

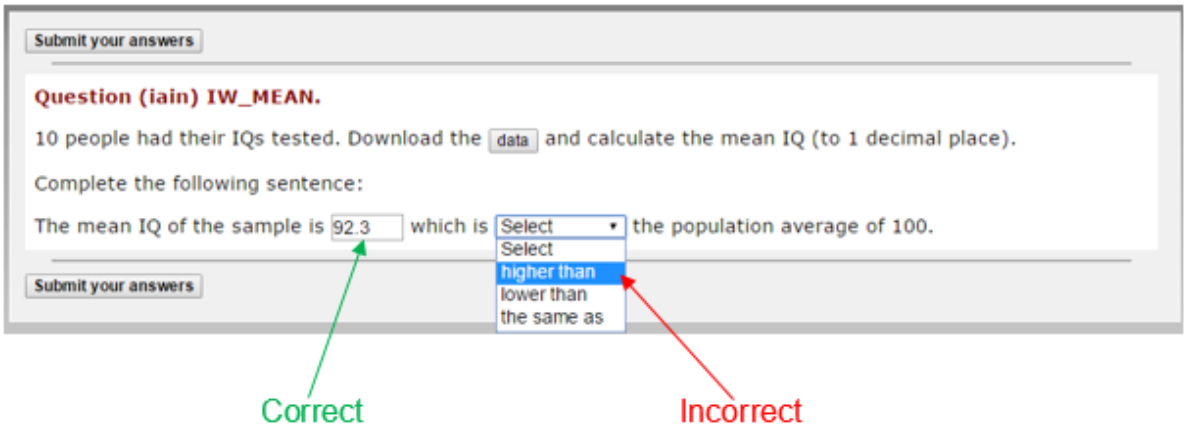


Figure 2. Screenshot of Example 1: Calculating and comparing a mean. For illustration purposes the student has entered a correct mean but has an incorrect comparison.

Figure 3 demonstrates the feedback screen. The 'Question' section reminds the student of the question and the blue text indicates where numerical values or dropdown menu choices were required. The 'Solution' supplies the correct answers to the sentence the student needed to complete. The 'Report' indicates, with colour coded marking, what the student has answered

correctly or incorrectly. Incorrect answers can be supplemented with further formative feedback. In this example there is a sentence further explaining the students mistake.

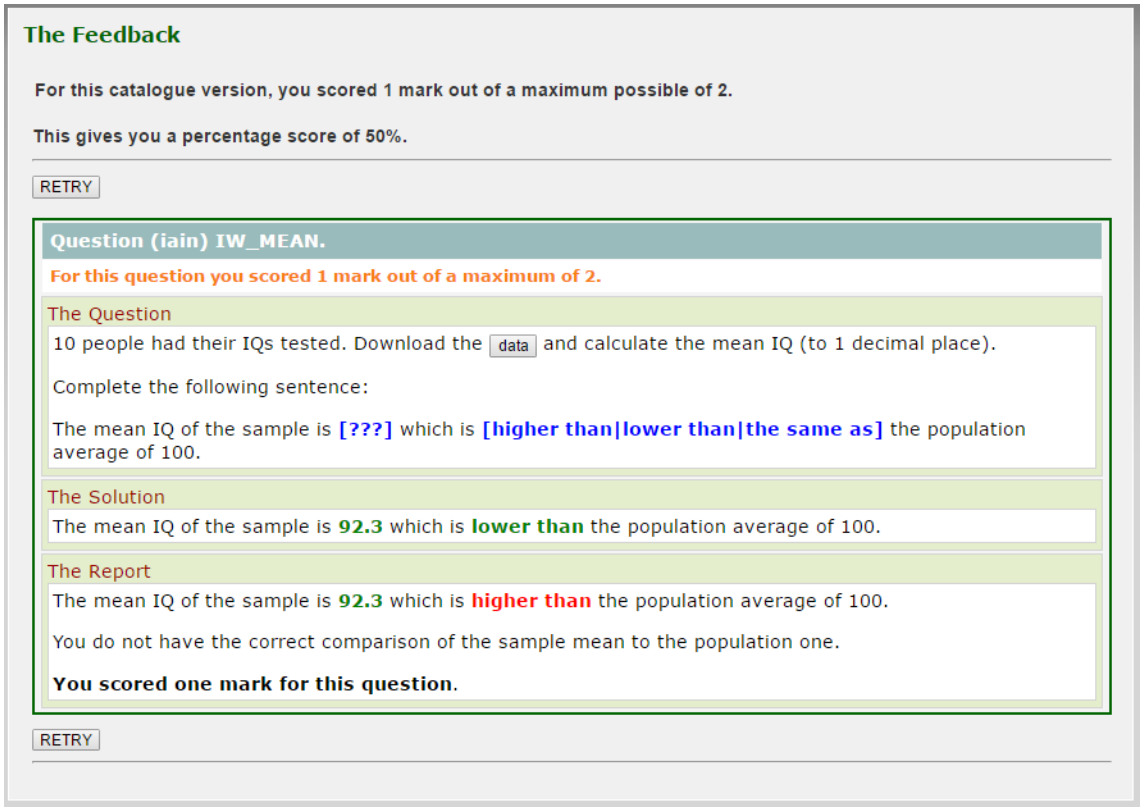


Figure 3. Screenshot of the feedback for Example 1 which comprises three sections: 'Question', 'Solution' and 'Report'. Note that the latter indicates that the student has an incorrect answer to the mean comparison.

The Dewis-R interface requires a script file that comprises two parts; the Dewis part relates to install-time question construction and the R part relates to run-time data generation and analysis. We shall now present the script that generates Example 1. Note that the explanation of syntax in any depth is beyond the scope of this article. The syntax we are developing is in its infancy; in time, our aim is to publish a manual with full details.

Figure 4 presents the Dewis install-time question communication script; it comprises HTML code and our own coding tags which we shall briefly explain. The <PARAMETERS> tag code is where the number and type of each question input is defined, each of which is assigned an ID number; here we have two. The <INPUT> tags define the first (ID=1) to be numerical and the second (ID=2) to be a dropdown with three choices. The name of the R run-time function that will generate data and input answers are also defined. The <ON_SCREEN> tag code is where the question text is composed. Question and feedback reporting are in a format that has three sections; 'Question', 'Solution' and 'Report'. The <SHOW_IN> and <HIDE_IN> tags dictate what text appears in each of the sections. <IF_WRONG> tags allow appropriate feedback relating to any wrong answers.

```
##########################################################################
################## DEWIS install-time question construction ##########
##########################################################################

<DEWIS_INSTALL>

#====================== Define run-time function and define inputs ===

<PARAMETERS>

 <RUN_TIME_FUNCTION='dewis_run()'/>

 <NUM_IDS=2>

 <INPUT TYPE=NUMERICAL ID=1 NAME='the mean IQ'/>

 <INPUT TYPE=DROPDOWN ID=2 NAME='the mean comparison'>
  <OPTION>higher than</OPTION>
  <OPTION>lower than</OPTION>
  <OPTION>the same as</OPTION>
 </INPUT>

</PARAMETERS>


#====================== On screen question text ======================

<ON_SCREEN>

<SHOW_IN='QUESTION'>
 10 people had their IQs tested. Download the <DATALINK>data</DATALINK> and
 calculate the mean IQ (to 1 decimal place).
 <p></p>
 Complete the following sentence:
 <p></p>

<SHOW_IN='REPORT SOLUTION'>
 The mean IQ of the sample is <INPUT ID=1/> which is <INPUT ID=2/> the
 population average of 100.

<HIDE_IN='QUESTION SOLUTION'>
<IF_WRONG ID='1'>
 <p></p>
 You have not supplied the correct sample mean value.
</IF_WRONG>
<IF_WRONG ID='2'>
 <p></p>
 You do not have the correct comparison of the sample mean to the population one.
</IF_WRONG>

</ON_SCREEN>
</DEWIS_INSTALL>
```

Figure 4. Example 1 script: Dewis install-time question construction.

Figure 5 presents the script that comprises the R function that is called at run-time. This function deals with all data generation and calculations plus a written communication of the correct input answers and data values for Dewis to read.

```
#<R>

##########################################################################
################## R run-time function ###################################
##########################################################################

dewis_run=function(){

#====================== Data generation and calculations =============

# get mean from 10 integer observations from N(100,15^2)
 IQ=round(rnorm(10,100,15))
 meanIQ=round(mean(IQ),1)
 IQdata=data.frame(IQ)

# comparison to mean of 100 1="higher"/2="lower"/3="the same"
 if (meanIQ>100){
  meanComp=1
 }else if(meanIQ<100){
  meanComp=2
 }else{
  meanComp=3
 }


#====================== Assign and communicate correct answers =======

 cat("
 <DEWIS_PARAMS>

# Assign correct answers for each input
 <CORRECT ID=1>",meanIQ,"</CORRECT>
 <CORRECT ID=2>",meanComp,"</CORRECT>

# Printing out of generated data
 <DEWIS_DATA COLS=1>
 ")
 print(IQdata)
 cat("
 </DEWIS_DATA>

 </DEWIS_PARAMS>
 ")

}
```

Figure 5. Example 1 script: R run-time function.

2.2. Example 2: Reporting a correlation

The second example is more adventurous and relates to reporting a test of Pearson's correlation coefficient. We shall restrict ourselves merely to screenshots that demonstrate some of the other tags that we are developing. Figure 6 is a screenshot of the question with various tags identified. In this example <DATALINK> provides a button for downloading the bivariate data of random sample size that has a correlation that is of a random strength and sign. The file format in this case is a CSV file, which is appropriate considering the wide range of operating systems used by students. Dynamic graphics generated via R during run-time can be included in questions using the <IMAGE> tag; in this example a scatterplot of the data is displayed. Any HTML tags may be included in a question; here we use <a> to include a link to a supporting video. Depending upon a significant or non-significant test result, different reporting statements are required; the <AREA_CHOICE> tag allows the student to pick the appropriate statement to report. In figure 7 the two competing report statements are displayed; it can be seen that extra inputs are required when reporting a significant correlation.

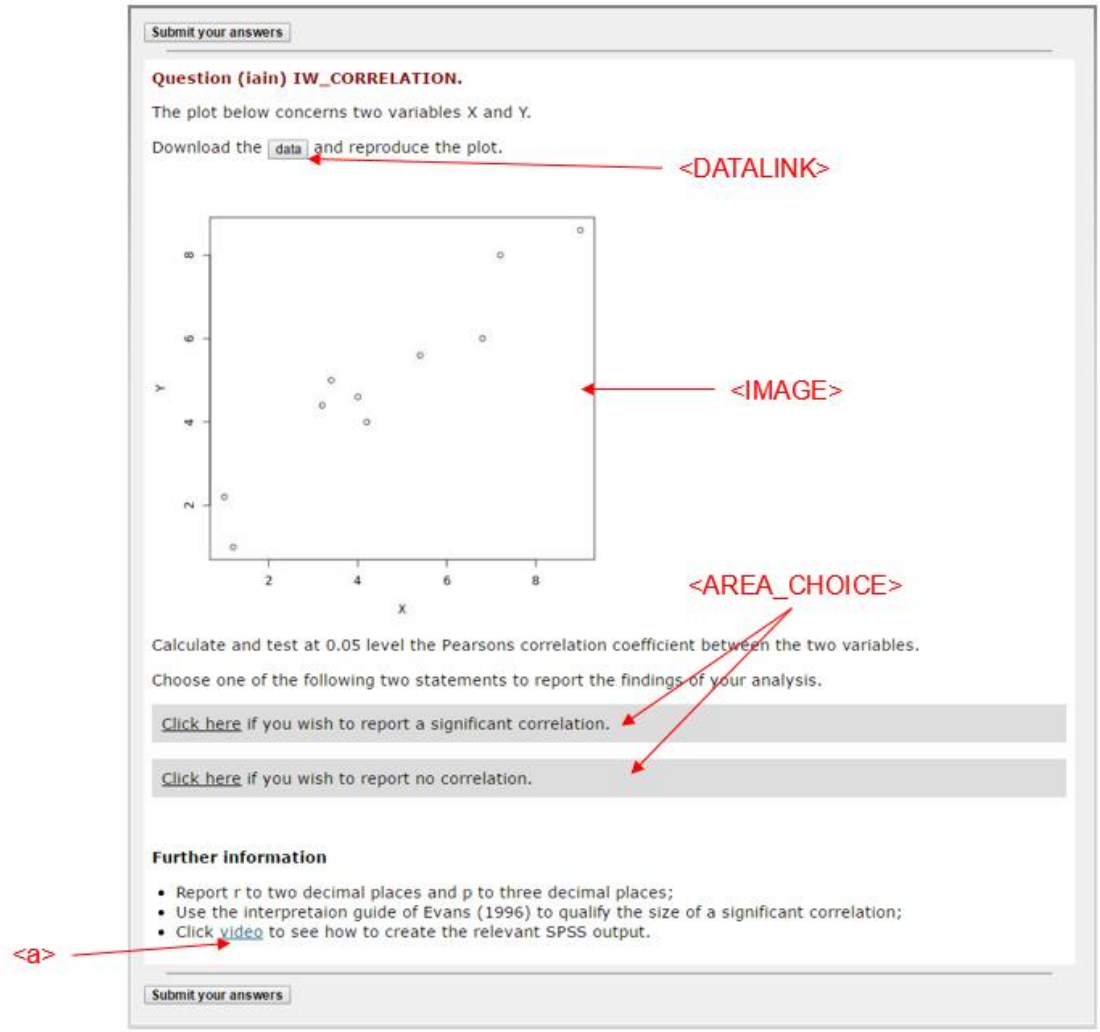
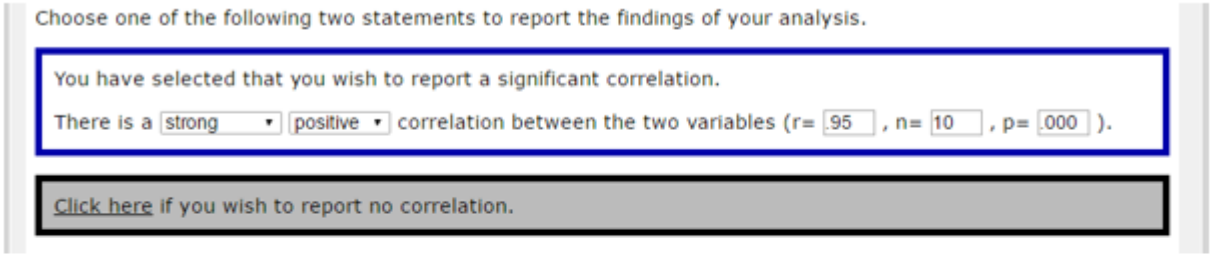


Figure 6. Screenshot of Example 2: Reporting a correlation. The use of various script tags is indicated.

a)



b)

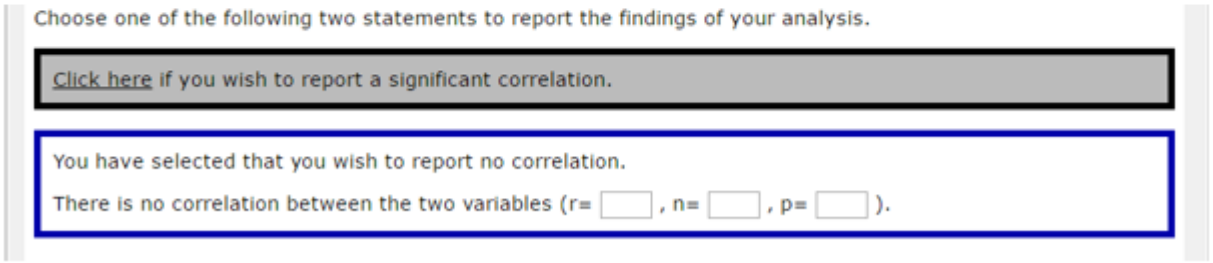


Figure 7. Example 2 competing correlation reporting statements. In a) additional information about the strength and sign of a significant correlation is required when reporting a significant correlation that is not necessary in b) when reporting that there is no correlation.

3. Discussion

Although early in the development stage, we have illustrated, through two examples, the power and ease with which meaningful statistics e-assessments can be created. This approach has allowed other statisticians within UWE to author questions and use e-assessment in their modules for the first time. We would welcome collaboration from the wider statistics community to trial this new approach of authoring and to become involved in the development of the interface functionality. We anticipate that the creation of a Dewis-R statistics e-assessment community will help the dissemination of this methodology and thus enhance the student experience. Staff in the community will benefit both from the sharing of learning & teaching innovation and the statistics practice aspects.

4. References

Dewis resources on **stats**tutor, 2015. Available at: http://www.statstutor.ac.uk/search/?q=dewis [Accessed 14 May 2016].

Dewis welcome page, 2012. Available at: http://dewis.uwe.ac.uk [Accessed 14 October 2016].

Dewis-R Statistics Resources website, 2016. Available at http://dewisprod.uwe.ac.uk/resources/stats/index.html [Accessed 19 October 2016].

Gwynllyw, D.Rh., Weir, I.S., and Henderson K.L, 2016. Using DEWIS and R for multi-staged statistics e-assessments. *Teaching Mathematics & Its Applications*, 35(1). pp.14-26.

R Core Team, 2014. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available at: http://www.R-project.org/ [Accessed 23 December 2016].

Sangwin, C., 2013. *Computer Aided Assessment of Mathematics*. Oxford University Press.

Sigma Resource Development Projects Funded, 2015. Available at: http://www.sigma-network.ac.uk/resource-development-projects-funded/ [Accessed 14 May 2016].

Weir, I., Gwynllyw, R. and Henderson, K., 2015. Using technology to inspire and enhance the learning of statistics in a large cohort of diverse ability. In: IATED, ed. (2015) *Edulearn15 Proceedings*. IATED.