Evaluation of a Questionnaire to Measure the Impact of Mathematics Masterclasses at the University of South Wales[1]

A. Little, D.J.J. Farnell, and M. Jones

Division of Mathematics and Statistics, Faculty of Advanced Technology,

University of South Wales, Pontypridd CF37 1DL, Wales, United Kingdom

Abstract

The Royal Institution began the mathematics masterclasses in order to foster mathematics outreach activities from academia to the wider community. Here we validate a questionnaire that has previously been used to measure the impact of these masterclasses held at the University of South Wales (formerly known as the University of Glamorgan) in 2010 and 2011. We administered questionnaires to 111 key-stage 3 (KS3) school children and to 76 key-stage 4 (KS4) school children recruited from local schools in the south Wales area. Cronbach's α coefficient was found to be greater than 0.7, thus indicating that the questionnaires were internally consistent. Exploratory factor analysis demonstrated that considerable structure occurred in the responses from both the KS3 and KS4 students. The questionnaire was broken down into five subsections or "subscales" (made up of related questions), namely: enthusiasm and (separately) confidence in mathematics; future plans of the students involving mathematics; course delivery and content; and, perceived usefulness of mathematics in society. Results for the confidence and usefulness subscales were found to be high for both student groups. A reduced level of enthusiasm (adequate only) was seen for the KS3 students, which might have been due to unrealistic expectations for the masterclasses prior to the course in this cohort of very young students or it might have been a characteristic of this particular cohort of students. An overall questionnaire score indicated that the masterclasses were generally well-received by the students, which makes sense because only the most mathematically able students were selected to participate by their schools, although individual items and subscales gave a more nuanced picture. All of these results indicate that

[1] Formerly known as The University of Glamorgan (1992 – 2013)

the questionnaire was a viable tool that gave sensible results for our cohort of students. Potential refinements to the questionnaire are suggested.

Keywords: Mathematics, masterclass, impact

1. Introduction

The Oxford dictionary (2011) defines the term ‘outreach’ as ‘an organisation’s involvement with or influence in the community’. Mathematical outreach is therefore defined as an organisation’s influence on the community’s attitudes towards and enthusiasm for Mathematics. Mathematics outreach events are organised for several reasons; to promote the subject itself, to generate a wider interest in the mathematical sciences and to encourage more people to study mathematics beyond the standard level required by statute. Generally the purpose of a mathematical outreach event is to broaden young peoples’ learning experience and introduce them to areas of mathematics that may not necessarily be covered in the national curriculum.

Year on year, more stimulating and interesting mathematical events are being held in all areas of the UK and in a range of accessible formats. Universities are one of the leading contributors in mathematics outreach. Of course, the foundation of their work relies on their collaboration with local schools, particularly secondary schools. The number of students taking advanced-level (i.e., “A”-level) mathematics in the UK dropped by 18% from 1985 to 2005 (JCQ, 2012), such that the overall number of students taking the course was approximately 6000 students down on those numbers taking the subject in the 1990s (Smith, 2004; Sainsbury, 2007). In recent years however, the number of students taking mathematics A-Level has begun to rise dramatically with a 46% rise in numbers from 2005 to 2012 (JCQ, 2013). The London Mathematical Society (1995) stated that students entering higher education lacked the necessary ability and skills for university-level mathematics. Examples of outreach throughout the UK highlight the on-going actions of academic institutions whose intention it is to produce inspired young mathematicians. For instance, The University of Liverpool has played a vital role in mathematics outreach with their creation of the *FunMaths Roadshow*. The first of these events was held in 1999 by the Liverpool Mathematical society. The roadshow consists of 350 mathematics activities and puzzles aimed at students from as young as primary school year one, right up to University undergraduates (Liverpool Mathematical Society, 2010). These activities are provided for students to attempt, usually as part of an organised outreach event. They can be attempted by students at their own pace and do not generally require any extra materials. This can be a great benefit as students can easily engage with the materials in an environment, which is somewhat different to that of a school classroom. The roadshow resources are now used in various locations throughout the UK and the overall aim is to encourage teamwork skills, problem solving and to stimulate discussion and interest in mathematics amongst students. A selection of the *FunMaths Roadshow* resources were used during one of the masterclass sessions entitled ‘puzzling mathematics’ held at the University of South Wales (formerly known as the University of Glamorgan).

Many schools and institutions have organised mathematics outreach events with the Royal Institution of Great Britain playing a leading role. The Royal Institution of Great Britain is a registered charity that organises science and mathematics events for young people in the UK. The aim of the charity is to inspire young people and spark their interest and enthusiasm for the scientific world. The charity was set up in 1825, starting with their well-known ‘Christmas Lectures’. These lectures are given once a year by an expert in the field; they have a different exciting theme each year and are now televised in the UK every Christmas. In 1978 Professor Sir Christopher Zeeman gave a Christmas lecture entitled ‘mathematics into pictures’, this lecture was very well received and inspired the idea for mathematics masterclasses, which commenced in 1981 (The Royal Institution of GB, 2011). These classes consist of a course of extra-curricular sessions for young people with an aptitude for mathematics. Similarly to the Christmas lectures, masterclasses will typically have a theme each week and will introduce an application of mathematics to the real world. Children are given opportunities to work in groups, discuss ideas and work on thought provoking puzzles and activities. Masterclasses are held in various locations all over the UK including the University of South Wales. Hayman (1989) discussed the initial issues that may have hindered the success of the masterclasses, such as few pupils being willing to participate or disinterested teachers. They go on to say that the classes have had some initial success over the years. This success was not initially reflected in the figures, which showed a drop in A-level entries in the 80’s. The “enrichment” and “engagement” aspects of “additional” mathematics teaching events such as (and in particular) the RI mathematics masterclasses were examined by Santos and Barmby (2010). Santos and Barmby (2010) recommend that masterclasses must encourage both mathematical activities to be carried out by the students themselves as well as engendering “mathematical thinking” in order to foster better engagement and enrichment. In this manner, Santos and Barmby (2010) state in their concluding statement that: “In engaging students through enrichment activities in mathematics, the possible benefit is the impact on the learning and understanding of mathematics, over and above the enjoyment and participation of students.” Feng (2010) reports in the abstract for their paper that: “Whilst students reported a range of enrichment benefits, broadly related to their mathematical, and personal and social, development, support for mathematics learning in school, and exposure to higher education, their experience was more subtly related to the characteristics of the programme in which they had participated, interpreted according to more familiar experiences of learning mathematics in school." Thus, additional mathematics "enrichment" activities (again such as the RI mathematics masterclasses) can only ever be effective if they are built on a firm foundation of a good "daily diet" of mathematics. Before we analyse the mathematics masterclasses at the University of South Wales, we will firstly review a similar study of other Institutions that hosted a RI masterclass series.

In 2008, the Curriculum, Evaluation and Management Centre at the University of Durham carried out an evaluation of the Royal Institution mathematics masterclasses. The study was based on various different masterclass sessions throughout the UK. A student feedback questionnaire was open to all students who had attended mathematics masterclasses and Barmby et al. (2008) summarised the results in their final report. Other studies were carried out within the investigation, such as questionnaires for teachers and organisers, but the student feedback is of most importance so we will review only this section of the report. (Note that the questions used in the questionnaire are used in our study and with the results presented in this article.) 971 students completed the questionnaire (445 males, 519 females and 7 did not state their gender). All students were aged between 12 and 14. One key finding from the Durham report was that the majority of students felt that they would like to study mathematics further than GCSE level (see Fig. 1) with 29% wanting to study for an undergraduate degree in mathematics. Although it is unclear of the contribution of the masterclasses to this result, it is a positive outcome and confirms that mathematics is being promoted effectively.

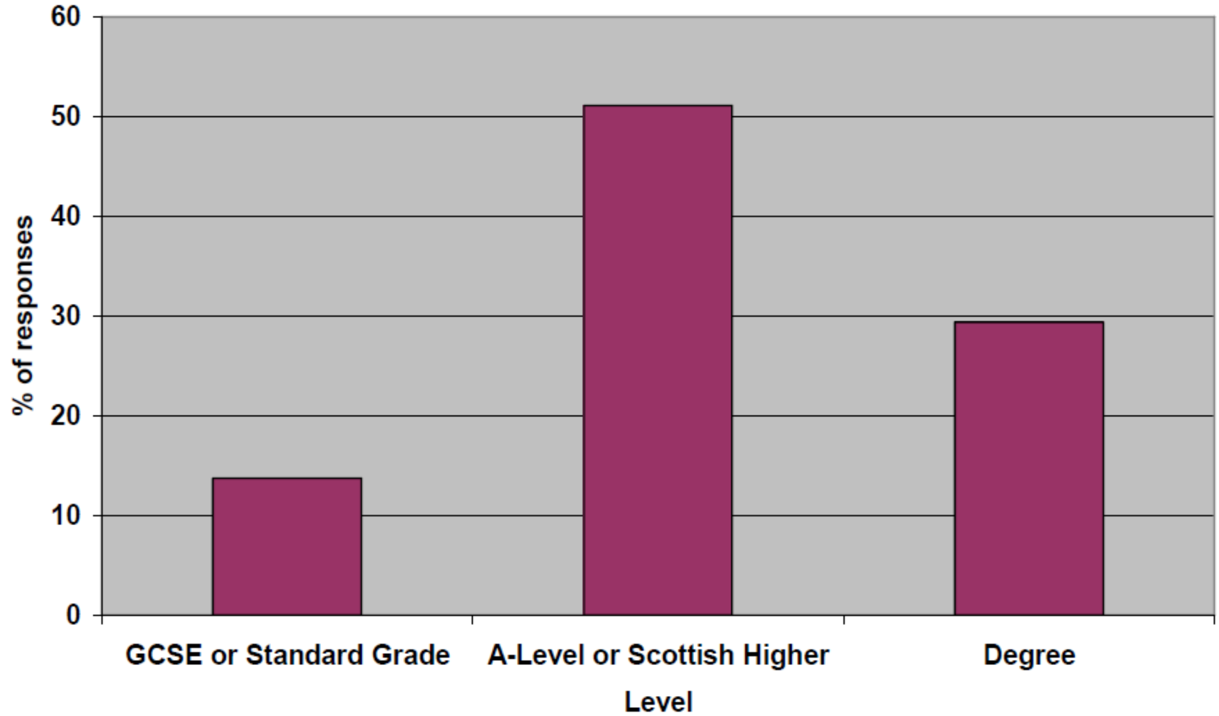

Figure 1: Graph summarising the level to which students felt they would study mathematics to, based on the results of the Durham University study (Barmby et al., 2008).

Questions relating to students feelings towards the course were given with a choice of responses from strongly disagree to agree and the results were summarised into a table quoting the percentage of students who agreed or strongly agreed with each statement. The statement that had the highest level of agreement from students was the statement 'I have had the opportunity to tackle new mathematical problems in the masterclasses', with 91.1% agreement. This shows that the masterclasses are achieving the objective of introducing young people to the type of mathematics which is different to that found in a usual classroom environment. The next two most agreed with statements were 'I enjoyed the classes' (84.3%) and 'I learnt a lot from the classes' (80%). This again supports the idea that the masterclasses are effective as the main aims are for students to learn new things and enjoy themselves at the same time. Interestingly, only two statements had an agreement percentage below 50%. The first being 'The masterclasses have encouraged me to pursue a career that uses mathematics' (39.4%) which is understandable as many students who enjoy mathematics may still not wish to use it in their career. The only other statement that had a low agreement percentage was 'the mathematics in the classes was too hard' (13.8%) which clearly is a positive outcome as most students did not feel too daunted by the difficulty of the mathematics.

Question number	Question	Percentage (%) of responses				
		Excellent	Very Good	Good	Poor	Very Poor
1	Quality of presenters and their presentations	30	45.9	21.5	1.4	0.0
2	Quality of the activities you carried out during classes	23.4	44.4	28.7	2.3	0.3
3	Quality of facilities where the classes were held	30.8	30.8	24.4	4.2	0.0

Table 1: Students opinions of the quality of the masterclasses quoted directly from the Durham report (Barmby et al., 2008). 971 students completed the questionnaire (445 males, 519 females and 7 did not state their gender).

Table 1 now presents results for the student's view regarding content and delivery of the masterclasses. The table shows that all aspects of the masterclasses were mostly scored as 'very good' and very few students voted any of the aspects as being 'poor' or 'very poor' which suggests that the general quality of masterclasses throughout the UK is very good. A set of questions was also asked based on activities students may like to get involved with in the future. The top response was 'mathematics trips and visits' with 38.2% of students saying they would be 'very interested'. The general response was that the overall aims and objectives of the masterclasses are being met and that the students are enjoying the classes; however there may be some issues such as timings and organisation that need to be addressed. The study carried out at Durham University was fairly extensive and covered various locations providing a wide range of information. The questionnaire effectively uses a mixture of quantitative and qualitative analysis has been used to give thorough results. This study gives a good analysis of how effective the course has been from the students' points of view.

In the original work of Barmby et al. (2008), no validation of the questionnaire by using statistical techniques was carried out. In this paper, we wish to evaluate this questionnaire as a measure the impact of RI mathematics masterclasses for a cohort of "mathophile" key-stage 3 and 4 students taken from the South Wales are of the United Kingdom by using standard methods commonly employed in statistics for this task. We aim to demonstrate that results from the questionnaire are sensible, and we will suggest places where the questionnaire can be refined. We will show that results of factor analysis have structure that is interpreted readily and that values for Cronbach's alpha coefficient (a measure of internal consistency) is "high" (i.e. it is near to a value of 1). We will show also that item, subscale, and overall questionnaire scores are towards the positive end of the spectrum, which is exactly what one would expect from the cohort of "mathophile" students and so again are "sensible". The methods section of this article describes how students were selected to attend the masterclasses series at the University of South Wales, as well as the characteristics of these students. We describe how the questionnaire was constructed and administered, as well as the statistical methods used. We present our results and then we proceed to discuss our conclusions in the final section of this article.

2. Materials and methods

2a. Sample Selection

The sample selected for this study was a group of secondary school students that attended the mathematics masterclass sessions at the University of South Wales during the academic year 2010/2011. All of those students participating in the masterclass series were selected because they had demonstrated to their teacher, an aptitude for and interest in mathematics. Teachers were specifically asked: "nominees should be mathematically able students who show promise and whom you feel would welcome being more fully extended mathematically." The students were handed an optional questionnaire at the end of the course. Both Key Stage 3 students aged 12-14 and Key Stage 4 students aged 14-16 were recruited as part of the study. 111 responses were received from KS3 (52 male and 59 female). From KS4, 76 responses were received (27 male and 49 female). All items in the questionnaire were completed by the students, and so "complete-case" analyses could be carried out.

2b. Questionnaire Design

In order to be consistent with earlier studies presented in the paper by Barmy *et al*. (2008), the questions used in the questionnaire are those items/questions in this paper. Indeed, the purpose of our paper is to validate this questionnaire for our student cohorts. Hence, it makes sense to go into some detail about the construction of this questionnaire. Each questionnaire consisted of a variety of statements relating to the students' attitudes towards mathematics. Some examples of these typical items in the questionnaire included: 'mathematics is a useful subject'; 'I am confident when solving mathematical problems'; and, 'I would like to study mathematics at A Level'. The students were asked to score their level of agreement with each statement according to the following scale: 1 = strongly disagree; 2 = disagree; 3 = neither agree nor disagree; 4 = agree; and, 5 = strongly agree. One question, which appears on all the questionnaires (KS3 and KS4, both prior to and after the masterclasses), reads 'mathematics lessons are boring'. We note that this particular item reversed the scale because 5 represented a negative response and 1 represented a positive response. The data for this item has been altered so that the question now reads 'Mathematics lessons are NOT boring' and the responses were reversed numerically (i.e., 6 – item value) so that all items were measured on a common (positive) scale. (Lists of the items used in the questionnaire are presented in tables below.)

2c. List of Lectures Undertaken during the Masterclass Series in years 2010 to 2011

In total 14 were lectures were given during the series of RI masterclass series at the University of South Wales in years 2010 to 2011. The list of titles for the KS3 and KS4 masterclasses in 2010/2011 were:

KS3 – "Puzzling Mathematics"; "Population Mathematics"; "Juggling Mathematics"; "The Mathematics of Google and the IPod"; "Artistic Mathematics"; "Sporting Mathematics" "Origami Mathematics"

KS4 – "More Puzzling Mathematics"; "Chaotic Mathematics"; "Sedoc Terces"; "Einsteinian Mathematics"; "Intercept Mathematics"; "Learning from Data"; "The Magic of Computation"

2d. Statistical Analyses

In order to present a general idea of the students' responses to the questionnaires, some simple descriptive statistics have been calculated and quoted with the results. Both mean and median scores for items and various subscales were determined.

2e. Subscale Construction

Subscales are collections of similar items into a class. We select the "similar items" in each scale "heuristically" (i.e., by using common sense). Using this common-sense approach, we believe that there are five subscales, although we do also validate this by using exploratory factor analysis (explained below). The usefulness of subscales is therefore that they measure some underlying factor much better than any single item. The subscales used here were:

Enthusiasm – questions relating to personal enthusiasm for and enjoyment of mathematics

Confidence – questions relating to confidence to use mathematics

Future – questions relating to plans to use and be involved in mathematics in future

Course – questions relating to the content and delivery of the masterclass course of lectures

Usefulness – questions relating to the perceived usefulness of masterclasses in society

Each statement in the questionnaire was assigned to a subscale and analysed accordingly. Note that an overall questionnaire score could be obtained with respect to all items and subjects. Mean and median item, subscale, and questionnaire scores were found.

2f. Exploratory Factor Analysis

There are broadly two types of factor analysis, namely: confirmatory factor analysis, which uses maximum likelihood methods; and, exploratory factor analysis, which uses principal component analysis (PCA). Confirmatory factor analysis tends to be used to confirm some preconceived structure between the items suspected or known prior to analysis, whereas exploratory factor analysis is often used for the initial to explore the relationships between items in the questionnaire. Exploratory factor analysis is probably the more commonly used techniques of factor analysis and here it was carried out using SPSS (Armstrong & Soelberg 1968, Brace et al., 2003, Field, 2005, Howitt, Cramer, 2008, Gorsuch, 1974).

Exploratory factor analysis carries out PCA of a (Pearson) correlation matrix in order to find both the eigenvalues λ_i and eigenvalues u_i of this matrix. As mentioned in the book by Field (2005), the eigenvalues λ_i provides an indication of the importance of each factor in explaining overall "variability" (of all results in the questionnaire) and that "the basic idea is that we retain factors with relatively large eigenvalues, and ignore those with relative small eigenvalues." Indeed, a commonly used rule called the Kaiser criterion is the ''eigenvalues greater than 1 rule" to establish the number important factors (i.e., the number of subscales). Factors f_i are vectors that are obtained from the eigenvalues (after rotation of axes, see below) by using: $f_i = u_i \times \sqrt{\lambda_i}$. This approach has the advantage that individual elements of the factor (vector), which are called factor loadings, may now also be used to explain the importance of specific items in specific factors to the overall correlation between items observed the questionnaire. This is a strong advantage of this method. The cut-off point for a factor loading to be considered "important" in this way is taken (here) to be 0.5.

It is well-known that principal components analysis of a covariance matrix for a set of correlated normally distributed variables is known to provide the principal axes (e.g., for a set of Cartesian coordinates $\{(x,y)\}$) because the covariance matrix provides information not only about correlation between the variables but also about their variances. However, correlation matrices do not contain any information relating to variance, and so eigenvalues from PCA (although still orthogonal) based on the correlation matrix may not lie along the principal axes. It is integral to application of exploratory factor analysis (Field, 2005) to carry out additional unitary transformations (i.e., rotations of the axes of the variables) after PCA in order to orient u_i along these principal axes. Note that unitary transformations do not affect eigenvalues and so we retain the important property that eigenvalues λ_i still yield an indication of the importance of each factor in explaining overall correlation between items. Again, factor loadings are found (now after rotation of axes in u_i) using: $f_i = u_i \times \sqrt{\lambda_i}$. There are two methods of carrying out the rotation of axes, namely, orthogonal and oblique. Orthogonal rotations assume that principal axes are indeed orthogonal, whereas oblique rotations do not make this assumption. Here we use Varimax rotation, which is a form of orthogonal rotation that also maximises differences between eigenvalues after rotation.

Again, we remember that the cut-off point for a factor loading to be considered important was taken to be 0.5. Questions with small factor loadings (<0.5) were omitted from the table of factor loadings for the sake of clarity. Note that the component values *in the columns* for each factor indicate groups of highly correlated questions. Hence, those items with high factor loadings indicate the items in the possible subscales. This process allowed us to check the construction of our proposed subscales given above. The Kaiser-Meyer-Olkin measure gives us an idea of how effective factor analyses has been (Field, 2005) and this was generally very high (KMO scores were generally > 0.7); such high values indicate that factor analysis was a suitable method to apply to these data sets.

The use of factor loadings from exploratory factor analysis as "weightings" in forming subscale scores is inappropriate because high factor loadings in the same factor are indicative of correlations between these variables only (they cannot be interpreted in the same way as, say, regression coefficients). The logical and standard approach is to use exploratory factor analysis to explore and inform the creation of subscales and to make weightings of items to the subscales equal, as described above. Factor analysis is used here to validate the questionnaire by determining if it has readily interpretable structure, and (secondly) to validate construction of subscales. Factor analysis can also be used to refine questionnaires, e.g., by condensing those items that have extremely high correlations into one single item. For example, this process was carried out in references by Ho *et al.,* (2009) and Farnell *et al.*, (2010) to reduce the size of questionnaires that measured patient-reported toxicity levels after radical cancer treatment for use in a busy clinic.

2g. Cronbach's α Coefficient

As mentioned in the book by Field (2005), it is often useful to find Cronbach's α when one wishes to carry out factor analysis. Cronbach's α estimates the association between items in a questionnaire by using a split-half methodology; the sum of one half of the items in the questionnaire or subscale is correlated with the sum of the other half of the items. However, we note that there are many ways of splitting the questionnaire into two halves and Cronbach's α approximates the average correlation between sums across all possible splits (without actually having to carry these "splits" explicitly). If there are *k* variables $\{x_1, x_2, \ldots, x_k\}$ either for all items in the questionnaire or for a specific subscale then Cronbach's is defined by:

$$\alpha = \frac{\boldsymbol{k}}{\boldsymbol{k}-1}\left(1-\frac{\sum_{\boldsymbol{i}=1}^{\boldsymbol{k}}\mathrm{var}(\boldsymbol{x}_i)}{\mathrm{var}(\sum_{\boldsymbol{i}=1}^{\boldsymbol{k}}\boldsymbol{x}_i)}\right)$$

A high value for Cronbach's α near to 1 shows that item scores are consistent with each other and a value near to zero shows that items scores are inconsistent. Cronbach's α is therefore often referred to as the "internal consistency" of a questionnaire. We note also that Cronbach's α is also sometimes referred to as a measure of "reliability" in the literature (e.g., in Field, 2005). However, we remark that Cronbach's α actually yields a lower bound to the "true statistical reliability," as is proven mathematically in the book by Dunn (1989). Field (2005) recommends that Cronbach's α should be found be found for all items if an overall if one forms an overall questionnaire score, and additionally separately for the subscales if subscale scores are found. A value for Cronbach's α greater than approximately 0.7 is taken to mean that the questionnaire or subscale is "internally consistent".

3. Results

3a. Factor Analysis

Factor analysis was carried out on both the KS3 and KS4 data taken using the questionnaires. We found that the Kaiser-Meyer-Olkin statistic was generally found to be high (e.g., KMO = 0.731 for the results in Table 2) and so the factor analysis ought to be "valid". (Note that a low score for the KMO value, i.e., near to zero indicates that the factor analysis was "invalid".) Generally between 4 and 6 factors had eigenvalues greater than 1, as shown in the Scree diagram of eigenvalues plotted in Fig. 2 with respect to factor/component number for KS3 student data. This indicates that there were generally four to six strong factors accounting for variation in the answers provided by the students for the questionnaires. This result agrees with our decision to condense the information provided by the items in the questionnaire into 5 subscales, namely, "usefulness", "course", "future", "enthusiasm", and "confidence."

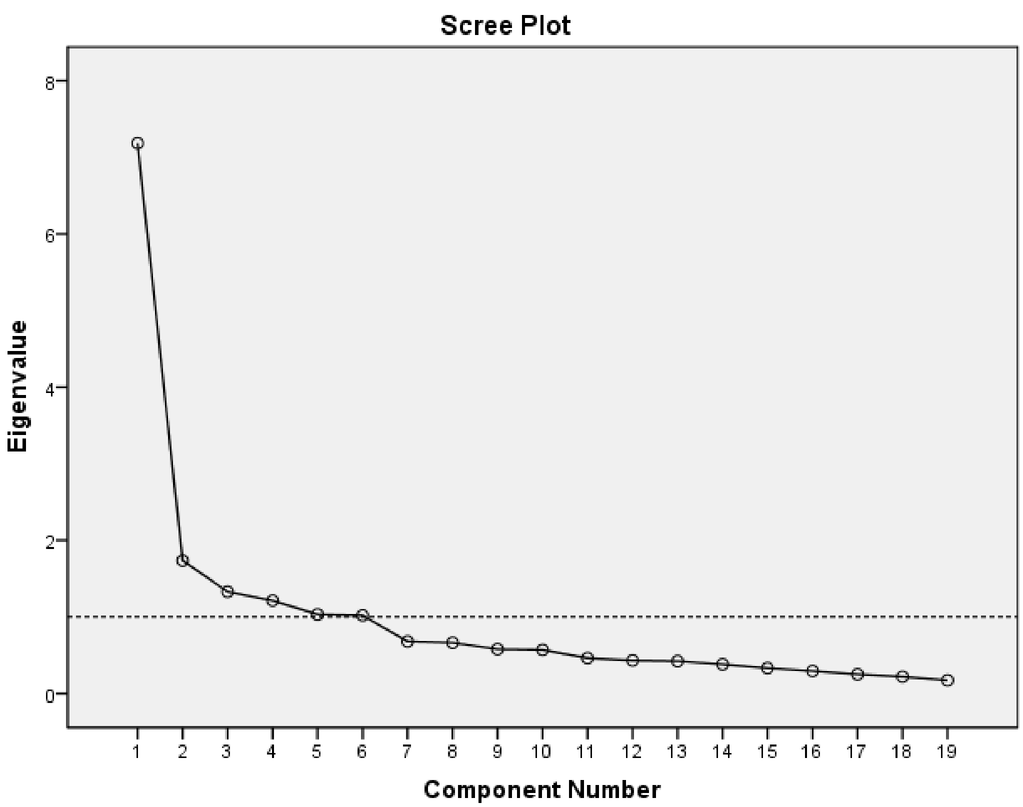

Figure 2: A typical scree plot of eigenvalues against factor number for the KS3 students.

Table 2 now shows typical results for exploratory factor analysis for the student feedback questionnaire (via SPSS v18.0 with VARIMAX rotation). Those factor scores that were less than 0.5 were omitted from the table in order to enable an analysis of the factor scores more readily. We see broadly that factor scores that were high in each column generally tend to stay within the proposed subscales, namely, "usefulness", "course", "future", "enthusiasm", and "confidence" (e.g., high factor loadings for the column for factor 2 relate exclusively to items/rows corresponding to "future" plans involving mathematics). However, strong correlations were also seen between items relating to the "course" and "confidence" for the KS4 data here (i.e., shown by high factor loadings for factor 1 relating to items/rows in both categories). Note that this was not seen for all of the other factor analyses carried out by us and so this result is just an anomaly. Again, all of these results provide excellent supporting evidence that our decision to divide the questionnaire into these subscales was a valid choice. Finally, it should be noted that creation of subscales from individual items is just a tool only (albeit a useful one), and that there is considerable freedom for constructing subscales from the items in questionnaires. The results of factor analysis (used in conjunction with common sense) are a useful guide for the creation of these subscales – or at least an independent check of them as part of the evaluation of the questionnaire tool. As part of the evaluation of the questionnaire, factor analysis strongly indicate that results for the KS3 and KS4 students are logical and reasonable, and so this is a useful test of the questionnaire.

Subscale	Question	1	2	3	4
Usefulness	Mathematics is a useful subject				0.732
	You need mathematics to understand the world				
	Mathematics is used extensively in today's world				0.684
Course	The masterclasses have helped me to see how useful mathematics is				
	Following the masterclasses I now have a better appreciation of mathematics	0.737			
	The masterclasses have helped increase my level of interest in mathematics.	0.544			
	The masterclasses have improved my skills in mathematics	0.731			
	I have learnt a lot of new things about mathematics at the masterclasses	0.709			
	I enjoyed the social side of the masterclasses	0.693			
	I enjoyed spending time with people who are enthusiastic about mathematics	0.510		0.554	
Future	I plan to study mathematics after GCSE		0.742		
	I plan to study mathematics at A-Level		0.849		

	I plan to study mathematics (or a related subject) at university		0.870	
	I will consider a career that uses mathematics		0.658	
Enthusiasm	Mathematics is one of my favourite subjects			0.638
	People who teach mathematics are interesting			0.592
	I enjoy solving mathematical problems			
	Mathematics lessons are not boring			
Confidence	I am generally confident in using mathematics	0.806		
	I am confident when solving mathematical problems	0.781		

Table 2: Typical factor analysis table for KS4 data. (KMO=0.731). (Factor loadings less than 0.5 are omitted for the sake of clarity.)

3b. Questionnaire Internal consistency

The questionnaires were found to be internally consistent by finding the Cronbach's α for all items in the questionnaire, which was generally found to be high (i.e., α=0.89 for the KS3 data and α=0.83 for the KS3 data). Results for Cronbach's α for the subscales were also high for both the KS3 and KS4 data (namely, KS3: Usefulness α=0.76, Course α=0.89, Future α=0.77, Enthusiasm α=0.64, Confidence α=0.73; KS4: Usefulness α=0.69, Course α=0.81, Future α=0.83, Enthusiasm α=0.52, Confidence α=0.84).

3c. Item and Subscale Results of the 2010/2011 Masterclass Questionnaire Data

KS3 Data

As part of the evaluation of the questionnaire, we now wish to consider average item and subscale results in order to determine if they make any sense. Taking the data as a whole, the KS3 data had a mean questionnaire score of 3.85 (evaluated over all items in the questionnaire) and the median was 3.84 (lower quartile=3.53; upper quartile=4.21). This indicates that in general the responses of the students were averaging nearer the positive end of the scale, so on the whole, a very positive response to the masterclasses. As one expects for our cohort of students, we conclude that the masterclasses had a broadly positive impact therefore.

	Question	**Means (Medians)**
Q1	Mathematics is a useful subject	4.72 (5)
Q2	The masterclasses will help/have helped me to see how useful mathematics is	4.28 (4)
Q3	I plan to study mathematics after GCSE	4.24 (4)

Q4	Following the masterclasses I now have a better appreciation of mathematics	4.05 (4)
Q5	The masterclasses have helped increase my level of interest in mathematics	3.89 (4)
Q6	Mathematics is one of my favourite subjects	3.98 (4)
Q7	The masterclasses improved my skills in mathematics	3.82 (4)
Q8	I am generally confident in using mathematics	3.86 (4)
Q9	I am confident when solving mathematical problems	4.35 (4)
Q10	I have learnt new things about mathematics at the masterclasses	3.65 (4)
Q11	People who teach mathematics are interesting	3.99 (4)
Q12	I enjoyed the social side of the masterclasses	3.66 (4)
Q13	I enjoyed spending time with people who are enthusiastic about mathematics	4.14 (4)
Q14	I plan to study mathematics at A-level	3.70 (4)
Q15	I plan to study mathematics (or a related subject) at university	4.09 (4)
Q16	I will consider a career that uses mathematics	3.94 (4)
Q17	I enjoy solving mathematical problems	2.44 (2)
Q18	Mathematics lessons are NOT boring	1.80 (2)
Q19	You need mathematics to understand the world.	4.51 (5)

Table 3: Summary of results for Questions 1-19 of the 2010/2011 KS3 questionnaire. Values are given for the mean and median (in brackets) item scores.

Table 3 shows the results for questions 1 to 19 of the 2010/2011 KS3 questionnaire for the median and mean scores. We see that only questions on that had low scores (mean scores below 3) related broadly to enthusiasm (Q17 and Q18). The highest scores (mean scores above 4) were related to the perceived usefulness of mathematics in society (Q1, Q3, Q19), future plans involving mathematics (Q3, Q15), sociable aspects of the course (Q13), and confidence in mathematics (Q9). We see that broad patterns are emerging from the data for similar items, thus indicating that results for the questionnaire do indeed appear to make sense for the KS3 data. This again supports the decision to divide items into separate subscales, and further insight can be gained by considering results for questionnaire subscales (below).

Subscale	Means (Medians)
Usefulness	4.60 (5)
Course	3.93 (4)
Future	3.99 (4)
Enthusiasm	3.05 (3)

Confidence	4.11 (4)

Table 4: Summary of results for subscales based on averaging item scores for questions 1-19 of the 2010/2011 KS3 questionnaire. Values are given for the mean and median (in brackets) subscale scores.

Table 4 shows the results for subscales relating to "usefulness," "enthusiasm," "course," "future," and "confidence" based on questions in the 2010/2011 KS3 questionnaire. Median and mean scores are presented. These results show again that the items that demonstrated the highest ratings related to the perceived "usefulness" of mathematics. We see that items relating to the "course," "future" and "confidence" also scored highly. Once again, we see that enthusiasm scored the lowest results for the KS3 students. These results show that the impact of the masterclasses can vary across different the criteria of "usefulness," "enthusiasm," "course," "future," and "confidence," and that an overall questionnaire score can miss these subtleties relating to impact.

KS4 Data

Taking the data as a whole, the KS4 masterclass data has a mean value of 4.07 and the median was 4.11 (lower quartile=3.72; upper quartile=4.44). This indicates again that in general the responses of the students were averaging nearer the positive end of the scale. Again, as one expects for our cohort of students, we conclude that this is a very positive response. Indeed, mean values indicate a slightly more positive response to the masterclasses than the KS3 students.

	Question	**Means (Medians)**
Q1	Mathematics is a useful subject	4.63 (5)
Q2	I plan to study mathematics after GCSE	4.18 (4)
Q3	Following the masterclasses I now have a better appreciation of mathematics	3.97 (4)
Q4	The masterclasses have helped increase my level of interest in mathematics	4.01 (4)
Q5	Mathematics is one of my favourite subjects	3.99 (4)
Q6	The masterclasses improved my skills in mathematics	3.99 (4)
Q7	I am generally confident in using mathematics	4.36 (4)
Q8	I am confident when solving mathematical problems	4.26 (4)
Q9	I have learnt new things about mathematics at the masterclasses	4.53 (5)
Q10	People who teach mathematics are interesting	3.70 (4)
Q11	I enjoyed the social side of the masterclasses	3.88 (4)
Q12	I enjoyed spending time with people who are enthusiastic about mathematics	3.76 (4)
Q13	I plan to study mathematics at A-level	4.16 (4)

Q14	I plan to study mathematics (or a related subject) at university	3.57 (3)
Q15	I will consider a career that uses mathematics	3.66 (4)
Q16	I enjoy solving mathematical problems	4.17 (4)
Q17	Mathematics lessons are NOT boring	3.74 (4)
Q18	You need mathematics to understand the world.	4.18 (4)
Q19	Mathematics is used extensively in today's world	4.50 (5)

Table 5: Summary of results for Questions 1-19 of the 2010/2011 KS4 questionnaire. Values are given for the mean and median (in brackets) item scores.

Table 5 shows the results for questions 1 to 19 of the 2010/2011 KS4 questionnaire for the median and mean scores obtained from masterclass. The mean and median scores for all items are found to be on the positive end of the spectrum (i.e., they are all greater numerically than 3). Again, we see that items relating to "usefulness" of mathematics (Q1, Q19) appear to score highest. We see also that the students appeared to have learnt much from the course (shown by Q9) and again they scored highly for confidence in mathematics (Q7,Q8) and short-term plans involving mathematics (Q2 and Q13). Interestingly, the item relating to the longer-term plans of students to involve mathematics (Q13). Most notably, items relating broadly to "enthusiasm" for mathematics (Q16, Q17) were much higher for the KS4 students compared to the KS3 students. As for the KS3 data, we see that broad patterns are emerging from the data for similar items, thus indicating that results for the questionnaire do indeed appear to make sense also for the KS4 data.

Subscale	Means (Medians)
Usefulness	4.44 (4.5)
Course	3.92 (4)
Future	3.89 (4)
Enthusiasm	3.90 (4)
Confidence	4.31 (4)

Table 6: Summary of results for subscales based on averaging item scores for questions 1-19 of the 2010/2011 KS4 questionnaire. Values are given for the mean and median (in brackets) subscale scores.

Table 6 shows the results for subscales relating to "usefulness," "enthusiasm," "course," "future," and "confidence" based on questions in the 2010/2011 KS4 questionnaire. Median and mean scores are presented and all are near to a value of 4, which is again well towards the positive end of the spectrum. This indicates again that the masterclasses were generally well received, as expected for our cohort of students. This table again shows a strong result for the "confidence" and "usefulness" subscales. The result for subscale relating "enthusiasm" is much higher for the KS4 students compared to the KS3 students, thus perhaps indicating a higher level of maturity for this group. Again, we see that impact can vary across the criteria of usefulness," "enthusiasm," "course," "future," and "confidence."

3d. Refinement of the questionnaire

The refinement of the questionnaire to measure the impact of the mathematics masterclasses can be carried out based on the statistical results presented here and also by using "common sense." Clearly, there are some items in the questionnaire that are very similar. Exploratory factor analysis is particularly useful in this case because those items with mutually very high factor loadings in the same factor might be condensed into a single item or one of the items might be omitted from the questionnaire. Note that a similar process was carried out for questionnaires to measure patient-reported toxicity after cancer treatment in a busy clinic where patient compliance in filling out the questionnaire can be affected by the length of the questionnaire (see Ho *et. al.* 2009, and Farnell *et al.* 2010). Thus, specific suggestions where items might be removed or combined (if so desired) are listed below:

Usefulness: "Mathematics is a useful subject," "The masterclasses have helped me to see how useful mathematics is" or "Mathematics is used extensively in today's world" could be replaced by a single item such as "Mathematics is a useful subject."

Course: "Following the masterclasses I now have a better appreciation of mathematics" and "The masterclasses have helped increase my level of interest in mathematics" could be replaced by a single item such as "The masterclasses have increased my appreciation of mathematics."

Course: "I enjoyed the social side of the masterclasses" and "I enjoyed spending time with people who are enthusiastic about mathematics" could be replaced by a single item such as "I enjoyed spending time socially with people who are enthusiastic about mathematics."

Future: "I plan to study mathematics after GCSE", "I plan to study mathematics at A-Level", "I plan to study mathematics (or a related subject) at university" could be replaced by a single item such as "I plan to study mathematics in future:" with responses "After GSCE," "To A-Level," and "To Degree Level," and "Not Applicable."

Enthusiasm: "People who teach mathematics are interesting" and "Mathematics lessons are boring" could be replaced by a single item such as "Mathematics is a boring subject."

Confidence: “I am generally confident in using mathematics” and “I am confident when solving mathematical problems” and could be replaced by a single item such as “I am confident in using mathematics and in solving mathematical problems.”

Finally, many of the items are on a positive scale (i.e., roughly going from 1=bad to 5=good). It is known that using a positive or negative scale for all items leading to a “response set bias.” This form of bias is the tendency of a participant to given items a cursory reading and to score all items similarly because all items are on either a positive or negative scale. One might wish to include some items on a negative scale, e.g., changing “I am generally confident in using mathematics” to “I am NOT generally confident in using mathematics” to overcome this form of bias.

4. Conclusions

In this article we carried out an evaluation of a questionnaire tool to measure the impact of RI mathematics masterclasses by using exploratory factor analysis, measures of internal consistency, and item and subscale scores. Exploratory factor analysis of results for items in the questionnaire for KS3 and KS4 schoolchildren showed much structure that appeared sensible. In particular, a common-sense division of the 19 or 20 items in the questionnaire into 5 subscales given by “usefulness”, “course”, “future”, “enthusiasm”, and “confidence” was generally supported by factor analyses. This is an important first step in the evaluation of the questionnaire, namely, that the factor analysis had readily interpretable structure and it has been successfully passed. However, factor analysis (along with common sense) was useful suggesting places where items might be removed or combined into a single item. The questionnaires were found to be internally consistent by finding the Cronbach’s α, which was generally found to be high (i.e., $\alpha > 0.7$). This is therefore another successful check of the questionnaire.

As expected from these results, we saw also that items scores were therefore generally similar for those items in the same subscales. For example, items related to “usefulness” scored highly (independently) for both the KS3 and KS4 data. An overall result for this study is that the masterclasses were generally well received by KS3 and KS4 students, which is evinced firstly by mean questionnaire scores of KS3 = 3.85 and KS3 = 4.07 (based on averaging values from all items in the questionnaire and over all students). Note that we should interpret the questionnaire scores (and subscale scores) from 1 to 5 as: 1 = very bad; 3 = neither good nor bad; and, 5 = very good. Hence, we see that our overall scores are towards the positive end of the spectrum. This demonstrates that masterclasses were indeed generally well received by the students and that the impact of the masterclasses was broadly positive. As only the most mathematically able students were selected by their schools to attend the masterclass series, one might well expect them to have high questionnaire scores overall and so these results appear to be logical, which is another successful check.

Similar patterns for results for the subscales are seen for the KS3 and KS4 students. In particular, the highest subscale scores were seen for the “usefulness” (mean: KS3=4.60 and KS4=4.44) and “confidence” (mean: KS3=4.11 and KS4=4.31) subscales, whereas the “course” (mean: KS3=3.93 and KS4=3.92), “future” (KS3=3.99 and KS4=3.89), and enthusiasm (mean: KS3=3.05 and KS4=3.90) subscales are slightly lower for both sets of students. An interesting result is that the KS3 students demonstrated somewhat low levels of enthusiasm for mathematics (where a value of 3 indicates “adequate” only); the impact of the masterclasses was lower for this cohort of students based on this criterion. We speculate that this might indicate that this is due to the higher maturity of the KS4 students, i.e., the topics discussed or the structure of the classes are more suited to the slightly older students and there may be a need for sessions which are slightly less in depth, or have a more exciting structure to suit the younger students. More data needs to be collected in order to establish more firmly if lack of enthusiasm was in fact just an anomaly for this particular group of KS3 students in 2010/2011. Indeed, we plan to continue to collect data in future series of mathematics masterclasses at the University of South Wales. We note however that there is never one definitive way to create subscales in a questionnaire and (even after factor analysis) that there is considerable scope for creativity in constructing such subscales. In any case, these results show that the impact of the masterclasses is subtler than any insight afforded by a single overall measure of impact, say, by an overall questionnaire score. We believe that such subscales are useful tools in evaluating the impact of the masterclasses for these questionnaires.

Finally, the results for the item, subscale, and overall questionnaire scores again appear sensible, and so this is another successful qualitative test of the questionnaire as part of its evaluation. These results, coupled with those results of factor analysis and internal consistency, indicate that the questionnaire is functioning adequately. The overall message of all of these results is that the questionnaire has been shown to be a successfully validated tool for measuring impact of RI masterclasses at the University of South Wales. Potential refinements of the questionnaire were possible though, and suggestions where different items might be removed or condensed into a single item have been presented (if so desired). It is recommended that a mixture of items measured on both positive and negatively scales be used in order to discourage “response set” bias. Future studies will involve pre- to post-masterclass analyses, as well as longitudinal trends by collecting many years of data in future. This will be aided greatly in future by the continued collection of quantitative (and qualitative) data relating to “impact” of mathematics masterclasses at University of South Wales. In this way, the Royal Institution mathematics masterclass series held at University of South Wales can continue to improve and evolve into a thoroughly effective and successful outreach programme for secondary school students to experience the exciting world of the mathematical sciences.

References

Armstrong J.S., and Soelberg P. (1968) on the interpretation of factor analysis. *Psychol Bull* **70**, 361–4.

Barmby P., Jones K., Kokotsaki D., Francis Ndaji F. and Searle J. (2008) Evaluation: Royal Institution secondary Masterclasses Programme. Full Report, CEM Centre, Durham University.

Brace N., Kemp R. and Snelger R. (2003) SPSS for psychologists, 2nd Edition, Palgrave Macmillan, Basingstoke, UK. **(**Chapter 11: factor analysis)

Dunn G. (1989) Design And Analysis of Reliability Studies: The Statistical Evaluation of Measurement Errors. Oxford University Press.

Farnell D.J.J., Mandall P., Anandadas C., Routledge J., Burns M.P., Logue J.P., Wylie J.P., Swindell R., Livsey J.,West C.M.L. and Davidson S.E. (2010) Development of a Patient-Reported Questionnaire for Collecting Toxicity Data following Prostate Brachytherapy. *Radiotherapy and Oncology* **97(1)**, 136–142.

Feng W.Y. (2010), Students' Experience of Mathematics Enrichment. Joubert, M. and Andrews, P. (Eds.) *Proceedings of the British Congress for Mathematics Education*, 83–88.

Field A. (2005) Discovering Statistics Using SPSS, Sage Publications, London, Thousand Oaks, New Delhi. (Chapter 15: factor analysis.)

Gorsuch R.K. (1974) Factor analysis. Philadelphia, London, Toronto: W.B. Saunders Publ.

Hayman M. (1989) The Royal Institution Mathematics Masterclasses. *Teaching Mathematics and its Applications* **8**, 7–11.

Ho K., Farnell D.J.J., Routledge J.P., Burns M.P., Slevin N., Sykes A., and Davidson S.E. (2009) Developing a CTCAE patient questionnaire for late toxicity after head and neck radiotherapy, *European Journal of Cancer* **45**, 1992-1998.

Howitt D. and Cramer D. (2008) Introduction to Statistics in Psychology., 4th edition, Pearson Education Ltd, Harlow, UK. (Chapter 18: non-parametric tests; and chapter 30: factor analysis.)

Joint Council for Qualifications (JCQ, 2012), Gender breakdown of the number of entries to A-level examinations in sciences and mathematics 1985–2008**;** *ibid.* (JCQ, 2013), GCE entry trends.

London Mathematical Society / Institute of Mathematics and its Applications / Royal Statistical Society (1995). Tackling the mathematics problem. London: LMS.

Oxford Dictionaries (2011), Outreach**.** Available at: http://oxforddictionaries.com/definition/outreach

The Liverpool Mathematical Society (2010) FunMathematics Roadshow. Available at: http://www.mathematics.liv.ac.uk/lms/funmathematics/

Sainsbury D. (2007). The race to the top: A review of the government's science and innovation policies. London: HMSO.

Santos S., and Barmby P. (2010), Enrichment and engagement in mathematics. Joubert, M. and Andrews, P. (Eds.) *Proceedings of the British Congress for Mathematics Education*, 199–206.

Smith A. (2004). Making Mathematics Count: The Report of Professor Adrian Smith's Inquiry into Post-14 Mathematics Education. London: DfES.