5-YEAR FOLLOW UP SURVEY OF USAGE OF TOOLS GIVEN IN ONE-DAY SAMPLE SIZE COURSE

Angie Wade
Great Ormond Street Institute of Child Health
30 Guilford Street, London WC1N 1EH, UK
awade@ucl.ac.uk

Our one-day Introduction to Sample Size course has run 20 times since 2012. Comprehensive written notes are given for use in conjunction with custom made excel sheets to perform calculations and these are openly available via a weblink. Immediate feedback for the course is generally excellent. To gauge the extent to which course materials had ongoing usage, and whether post leaving the course perception of the material and its' usefulness remained, a short survey was emailed to all 383 course participants from the 20 courses. Attendees noted ongoing usage of the course materials, with over two thirds of respondents subsequently referring to these frequently or occasionally. Even the most conservative estimate showed a substantial proportion still gaining direct benefit after 5 years.

BACKGROUND

For 10 years the UCL Centre for Applied Statistics Courses (CASC) has presented 1 to 5 day courses primarily aimed at non-statisticians in the workplace. Attendees often aim to leave a course equipped to undertake their own research, as well as having an improved understanding of published research. The one-day introduction to Sample size estimation has proved highly popular and runs on average three or four times per year with an audience of up to 20 participants. Numbers are restricted by the fact that teaching takes place within a cluster room and participants are directed to custom made excel sheets to perform calculations. The Excel sheets were originally available via CDs given to students but are now accessed via a web link to which there is open access (http://www.ucl.ac.uk/ich/short-courses-events/about-stats-courses/samplesize). A folder of comprehensive course notes is given to each course participant detailing all course material and usage of the Excel sheets with examples.

All CASC courses request feedback from students via Opinio, a web-based tool used to gather ratings and comments. This feedback, collected shortly after the day of the course, is consistently positive. There remains the question of how useful the participants find the training and material on an ongoing basis. The sample size course is unique in that the Excel sheets are presented and their usage integral to the estimation processes introduced to the students. Furthermore, the Excel sheets remain available to the students for as long as they wish to use them. Therefore, the sample size course, with its very specific estimation tools as well as written notes, was ideal for beginning an investigation into the ongoing potential benefits of our training We decided to contact attendees to the sample size courses held over the last five years to evaluate whether they continued to use the Excel sheets and whether the course had in fact been of direct relevance to their work.

The aim of this survey was to address the research questions:

i) Did participants of the one day sample size course utilise the information and tools given in the course in their work life?
ii) Was usage related to time since attendance at the course?

DATA COLLECTION

An online survey was developed asking whether individuals had referred back to the course material and whether they had continued to use the Excel spreadsheets. Both of these questions had the response options: 'Yes frequently', 'Yes, occasionally', 'Rarely' and 'No'. The individuals were also asked whether what they had learnt on the sample size course had been of direct relevance and use to their work, with response options 'Yes' or 'No'. All questions had space for comments and there was an additional overall comment box at the end asking for any other suggestions that may help improve the one-day introduction to Sample size course. Potential

In M. A. Sorto, A. White, & L. Guyot (Eds.), *Looking back, looking forward. Proceedings of the Tenth International Conference on Teaching Statistics (ICOTS10, July, 2018), Kyoto, Japan.* Voorburg, The Netherlands: International Statistical Institute. iase-web.org

participants were sent an email requesting that they complete the web survey (via a link), which would consist of only 3 questions. By making the survey short and informing of length in the accompanying email, it was hoped to maximize the response rate (Galesic and Bosniak, 2009).

The survey was sent via email to all 383 participants of the 20 one day courses held between 2012 and April 2017. There were 46 emails (12%) returned undeliverable with the proportion falling over time from 17/74 (23%) attendees at the three courses in 2012 to only one of the 39 (2.5%) in the two courses in the first half of 2017. Since the survey was anonymous, and we did not wish to alienate future participants by repeated emails, no additional requests were made to those who may not have responded.

RESULTS

From the 337 emails that were successfully sent, responses were received from 62 (18.4%). Most of the respondents (39/62, 63%) completed the form in a minute or less and over 90% in less than 5 minutes (56), two outliers were 26 and 13 minutes. Over 80% (50) replied within 48 hours, and all but one within eight days. Response rates were not dependent on time since course attendance being between 16 and 20% for all but two years, which were 2014 (13.9%) and 2013 (25%).

Two people skipped the question about whether what they had learnt on the one-day Sample size course had been of direct relevance to their work. Of the 60 that submitted a completed survey form, the overwhelming response was ‘Yes’ (55/60, 91.7%) and this was consistent across years. The five who did not find the course relevant, rarely (3) or never (2) used the course material and also rarely (1) or never (4) returned to use the Excel sheets.

All individuals answered the usage questions. Most referred back to the course material and the excel spreadsheets occasionally or frequently. There was a strong correspondence between responses from the two questions (table 1). Two individuals occasionally referred back to the course material but did not use the Excel spreadsheets, a single individual made occasional use of the spreadsheets but did not refer to the course material.

Table 1: Ongoing usage of course material and Excel spreadsheets responses

	Have you returned to the excel spreadsheets and used these?				
	Yes, frequently	Yes, occasionally	Rarely	No	TOTAL
Have you referred back to the course material?					
Yes, frequently	7	1	0	0	8
Yes, occasionally	3	24	6	2	35
Rarely	0	3	6	6	15
No	0	1	0	3	4
TOTAL	10	29	12	11	62

When analysed according to year of attendance, there was no consistent pattern of more usage either if the course was more recent nor with increasing time since receiving (figure 1).

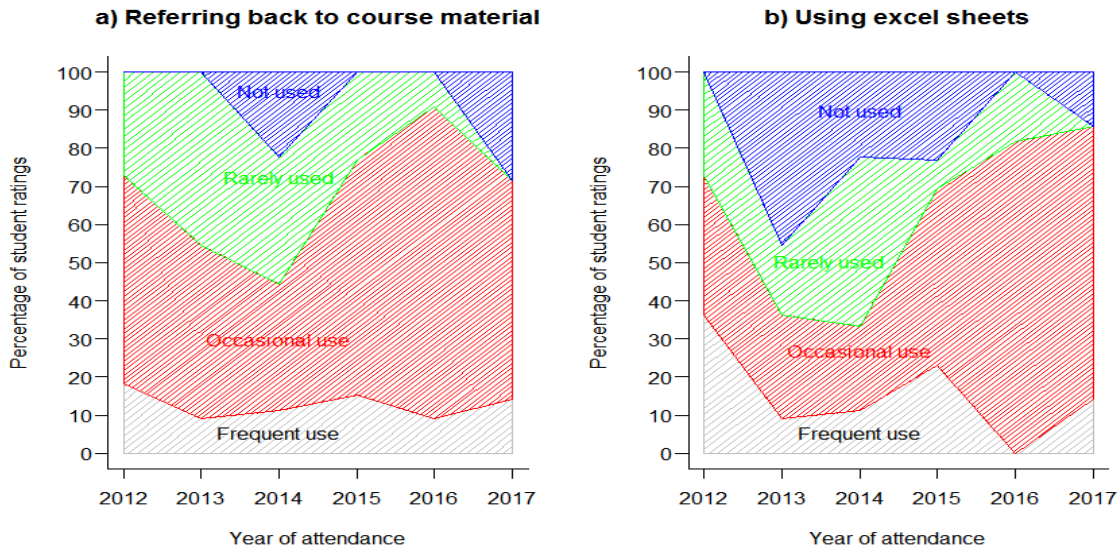

Figure 1: Temporal changes in responses to usage questions

Comments

A total of 53 comments were left by 34 individuals. Comments attached to the usage questions were most often given by those who responded 'Rarely' or 'No'.

There were six comments following the question about usage of course material with half from individuals who cited rare usage but explained their lack of need rather than problems with the material. One person who said they never referred back to the course material commented that they disliked using Excel, whereas another person who rarely used the course material stated difficulties following the concepts. The final comment was from someone who occasionally referred back to the course material *("2 colleagues have asked me to help them with a power calculation and one colleague borrowed the course file").*

There were 8 comments given following the question about usage of the spreadsheets, including the person who disliked Excel and reiterated their previous comment. Two of those commenting on this question stated that they had gone on to use Stata (occasional user) or G Power (never user). Two individuals said they occasionally used the Excel sheets and gave positive comments. The remaining three comments were very positive despite recording rare *("However, I think that this material it's very useful for working with"*) or no *("I think it will come useful when I need to calculate sample size"* and *"While I am very glad to have them and the process of reviewing them certainly helped understanding, the design of of* (stet) *research does require the calculations used the spread sheet"*)

Of the 10 comments left after the relevance and usefulness in work question, eight were from individuals who answered positively and the only minor criticism was that Stata may have been better to use on the course*("It was interesting but using more traditional software such as Stata would have been far more helpful."*). Two comments were from the five who said the course was not of direct relevance and use to their work. They said, "*I typically use a repeated measure design, which was not covered in the workshop*" and "*it will be useful when I get to design my own study and as I have been developing a proteomics method I have not needed to use the course info*".

27 of the 62 respondents left an overall comment at the end of the survey. Of these 16 made a comment that was purely positive with half of these specifically mentioning the usefulness of the Excel sheets. There was only one negative comment, from someone who attended the course in 2014, rarely used material or Excel sheet and did not find the course relevant to their work they said they "*did not learn anything as it was not explained*". However, a review of the opinio feedback from 2014 did not yield any such negativity. Three comments raised the issue of alternative software, with 2 of these also praising the Excel sheets. One user of GPower mentioned the need to quote software used for publication, *"I currently use G*power3 as it is more flexible*

*than the spreadsheets. Although the spreadsheets were useful I wonder if sometime spent on recognised software would be useful. I say this because, as part of getting published I have had to quote the software I used and the parameters I put in to generate the sample size and using G*Power3 the reviewers can check my work"*. Five comments were from individuals who would like more advanced topics covered: these were sample size calculations for repeat measures (2 comments), non-inferiority trials (these are given a brief mention in the course), unspecified 'epidemiological studies' and multiple groups (such as 5 ethnic groups). Despite each formula included in the course having at least one fully worked through example of usage, one participant wanted more. Their comment was *"more worked examples with answers would be useful when returning to the spreadsheets to be sure of using the correct one and inputting the correct figures"*. The final comment felt the course might be shorter, but seemed generally positive, *"was good, could be shorter, or just a half-day for more advanced users of excel/stats -overall it was very useful, if anything, it highlighted an often neglected issue in science "*.

DISCUSSION

A substantial proportion of those responding to the survey continued to refer to the course material and/or to use the online Excel spreadsheets for calculation. Some of those that did not currently use the material still cited the potential usefulness and intended to do so. Almost all felt that the course had been of direct relevance and useful to their work. There was no evidence of these findings being related to time since course attendance.

Strengths and Limitations

The response rate was low at 18.1%, but this is to be expected with a one shot survey of this type. Online survey response rates below 10% are not uncommon and rates are declining due to the volume that individuals currently receive (Van Mol, 2017). Those responding may have been biased towards those who did engage with the material. Conversely though they may have been biased towards those who found the course not very helpful. However, low response does not necessarily infer bias (Fosnacht et al, 2017).

We had access to email addresses of course attendees over the last five years. All presentations of the course were led by the same teacher and the content of the course remained similar throughout. We were asking about current usage and so we could evaluate the long term effect of the course. The numbers of non-deliverable emails declined over time as would be expected, but the response rates within those whose emails which were delivered did not show a temporal pattern.

If we consider the numbers responding positively from those where an email was delivered (ie. assume a worse-case scenario that those who did not respond rarely or never used the materials or spreadsheets) there remains 14% (95% ci (7.3, 25.3%)) of those who attended the course five years ago still gaining direct benefit occasionally or frequently from the course materials and the Excel sheets. We can consider this a very conservative lower limit.

CONCLUSION

This study gives evidence of a long term benefit of attendance at the one day sample size course, with over two thirds of respondents who had attended the course stating that they subsequently referred to the course materials frequently or occasionally. Even the most conservative estimate shows a substantial proportion of attendees continuing to use the course material directly in their work after 5 years.

REFERENCES

Fosnacht, K., Sarraf, S., Howe, E., & Peck, L. K. (2017). How Important are High Response Rates for College Surveys? *The Review of Higher Education*, *40*(2), 245-265.

Galesic, M. & Bosnjak, M. (2009) Effects of Questionnaire Length on Participation and Indicators of Response Quality in a Web Survey, *Public Opinion Quarterly*, *73*(2), 349–360.

Van Mol, C. (2017) Improving web survey efficiency: the impact of an extra reminder and reminder content on web survey response, *International Journal of Social Research Methodology*, *20*(4), 317-327.