Asking gender questions

Jonathan Pritchard, Karen Masters, James Allen, Filippo Contenta, Leo Huckvale, Stephen Wilkins and Alice Zocchi report on a survey of the gender of astronomers attending and asking questions at this year's UK National Astronomy Meeting.

Inspired by a recent report on the gender balance of astronomers asking questions at the 223rd Meeting of the American Astronomical Society (AAS) in Washington DC in January 2014 (Davenport *et al.* 2014), we decided to repeat the experiment at the most recent National Astronomy Meeting (NAM), held in Portsmouth in June 2014.

The gender balance of both speakers and session chairs at NAM (31% and 29% women respectively) closely matched that of attendees (28% female), but we find that women were under-represented among question askers (18% women). Women were especially under-represented in asking the first question (just 14% of first questions asked by women), but when the Q&A session reached four or more questions, women and men asked roughly equal numbers of questions. We found a small increase in the fraction of questions from women in sessions where the chair was also female (but this had no statistical significance). We find that on average 2.2 ± 0.1 questions were asked per talk, and observed no detectable difference in the number of questions asked of female and male speakers, but found that female chairs solicited slightly fewer questions on average than male chairs.

These results have some similarities to but also subtle differences from those reported by Davenport *et al.* (2014) for the AAS. They also found that the gender balance of speakers and chairs closely matched that of attendees (all roughly 35% women), and that women were under-represented among question askers (24% women). However, Davenport *et al.* found a significant effect that when a session chair was female, women asked more questions, and also found that women speakers were asked more questions (with an average of 3.2 ± 0.2 questions per female speaker) than their male counterparts (for whom the average was 2.6 ± 0.1). We note that Davenport *et al.* also report that more questions are asked on average at AAS talks than we found to be asked at NAM.

The National Astronomy Meeting is the larg-

est UK astronomical conference and provides a good opportunity to obtain representative statistics on the demographics and dynamics of the UK astronomy and geophysics community and their behaviour at a professional conference. The NAM 2014 participant list contained 624 names from universities and organizations primarily in the UK. Using an open source Python module (https://pypi.python.org/pypi/SexMachine) supplemented by manual classification, we obtained an estimate of the gender breakdown of NAM 2014 attendees based on first names, which results in 72% male (452), 28% female (172).

Data collection

The scientific programme at NAM was organized into eight plenary talks, a town hall meeting, and approximately 63 parallel sessions which together had 363 talks (for a review of the meeting see Bowler 2014). We used a modified version of the web form created by Davenport *et al.* to collect data on the gender of speakers, chairs and question askers in as many NAM talks as possible. The form asked for a talk ID (a concatenation of NAM session ID plus talk order was suggested), gender of the speaker, gender of the chair and then a string of letters representing the gender of people asking questions (e.g. FFM would be entered if three questions were asked, the first two by women, the last by a man). We added a free-form comment box (following the recommendation in Davenport *et al.*), but no comments were submitted. Data collection was volunteer-led, requested via
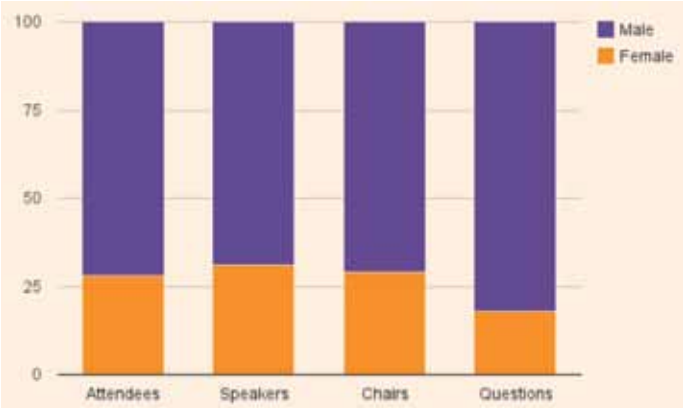
an email to participants from the NAM Local Organizing Committee before the conference began, and a daily Twitter campaign on the #rasnam2014 hashtag.

By the end of NAM this form had collected 595 separate submissions corresponding to 263 unique talks, which represented more than 70% of the scientific content at NAM. On average we collected 2.3 responses per talk. The completeness of this dataset is significantly higher than that analysed by Davenport *et al.*, who collected 300 responses comprising data on 225 talks, or about 26% of talks at the AAS meeting.
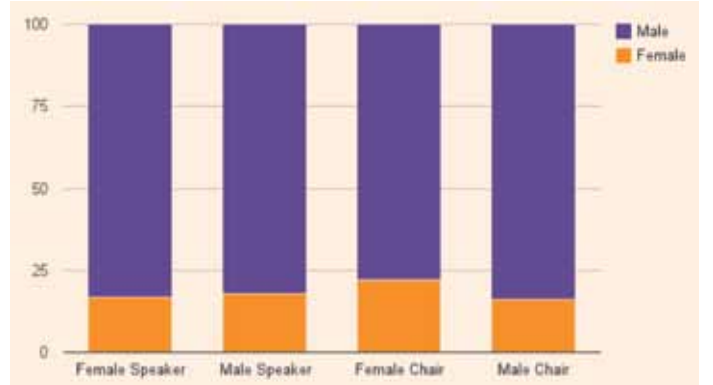
We began analysing the data during the NAM Hack Day (Simpson 2014). The raw data required extensive cleaning to correct for non-unique talk IDs, caused both by different naming of sessions and different numbering of talks within a session. Because we collected multiple answers per talk, we were also faced with examples where they did not agree.

Where a talk had more than one response, and there was no consistent majority answer, we retained longer question strings, assuming that shorter ones represented abbreviated or premature submissions of the form. We note that differences could also be due to ambiguity over an appropriate response, for example if the same questioner asks two questions at the same time, or if there was a back and forth between speaker and questioner. This was a factor for just a small fraction of entries: we do not expect it to influence our overall conclusions.

We report that at NAM 2014 the gender balance of speakers was $31 \pm 3\%$ women (we report

Table 1: NAM 2014 data

	women	men	total
attendees	172 (28±2%)	452 (72±2%)	624
speakers*	81 (31±3%)	181 (69±3%)	262
chairs*	75 (29±3%)	188 (71±3%)	262
question askers (QAs)*	101 (18±2%)	476 (82±2%)	577 (2.2±0.1 per talk)
QA female speaker	30 (17±3%)	150 (83±3%)	180 (2.2±0.1 per talk)
QA male speaker	71 (18±2%)	325 (82±2%)	396 (2.2±0.1 per talk)
QA female chair	33 (22±3%)	119 (78±3%)	152 (2.1±0.1 per talk)
QA male chair	68 (16±2%)	357 (84±2%)	425 (2.3±0.1 per talk)
asking 1st question	35 (14±2%)	216 (86±2%)	251
asking 2nd/3rd question	49 (19±3%)	224 (81±3%)	273
asking 4th–7th question	17 (32±6%)	36 (68±6%)	53

*These totals double-count people who chaired multiple talks (which is common, as most chairs are for a session of several talks), gave multiple talks, and those who asked multiple questions.

1: Summary of the gender of attendees, speakers, chairs and question askers at NAM2014.

2: Breakdown of gender of question askers by gender of speaker or chair.
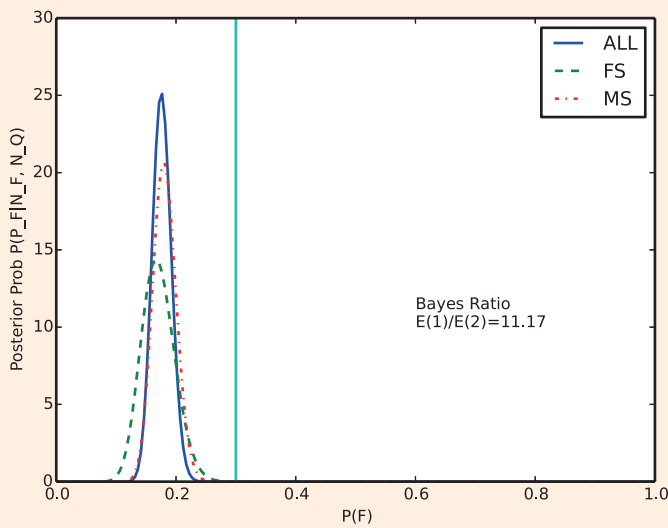
3: NAM 2014 data. Posterior probability of p_F, the probability that any individual questioner will be female. We compare the data for the full conference with that of subsets split by male or female speaker (left) or chair (right). The vertical line indicates the fraction of female attendees at NAM 2014. Also labelled is the Bayesian evidence ratio comparing our one- and two-parameter models.
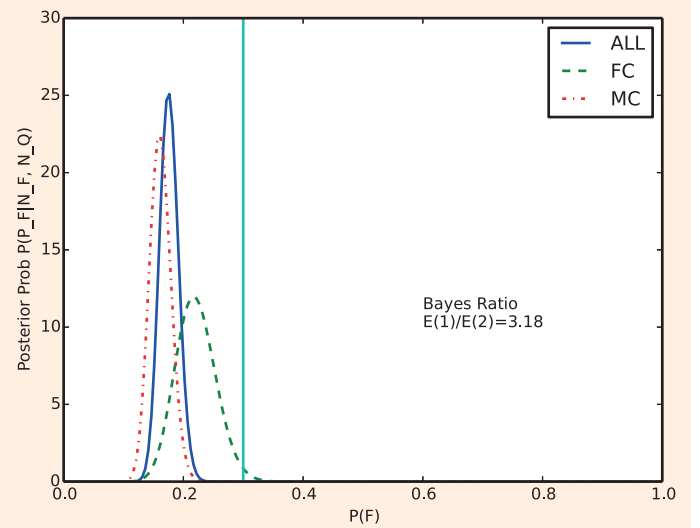
binomial uncertainties on these fractions) and that of chairs was $29 \pm 3\%$ women. This closely matched that of attendees, which we find to be made up of $28 \pm 2\%$ women. However, women were observed to be under-represented among question askers, which are just $18 \pm 2\%$ women. These data are shown in figure 1, and all our numeric data is summarized in table 1.

We define p_F as the probability that a given question will be asked by a woman. Under the assumption of binomial statistics for the gender of a question asker, an estimate of this probability can be found from the fraction of questions asked by women (i.e. we estimate $p_F = 0.18 \pm 0.02$). Obviously this value should equal the fraction of female attendees ($f_F = 0.28 \pm 0.02$) if questions are equally likely to be asked by men and women, which is not found to be the case.

From the data in table 1, we see that the mean number of questions per man was $\langle M \rangle = 1.05 \pm 0.03$ and per woman was $\langle F \rangle = 0.59 \pm 0.02$. If men and women asked questions equally these numbers would be the same, but clearly they are not. The ratio of these

numbers is $\langle M \rangle / \langle F \rangle = 1.79 \pm 0.06$ – a finding that male astronomers at NAM were roughly 1.8 times more likely to ask questions than female astronomers. Put another way, these data suggest that if NAM had had equal attendance by both men and women (i.e. if 50% of the attendees had been women), and in that circumstance the likelihood of a woman asking a given question remained the same as we have measured here, then almost two-thirds of questions would still be asked by male astronomers.

Impact of gender on questions

We break this down further, looking separately at the impact of the gender of both the speaker and the chair on the gender balance of question askers. We find no difference in the gender balance of question askers separated by gender of speaker (with $17 \pm 3\%$ and $18 \pm 2\%$ questions by women for female and male speakers respectively), and a small difference detectable with the gender of the chair ($22 \pm 3\%$ and $16 \pm 2\%$ questions were asked by women in talks chaired by women or men respectively). These data are shown in figure 2. This results in factors of

1.9 ± 0.1 and 1.7 ± 0.1 for how much more likely men are to ask questions than women of female and male speakers respectively; and factors of 1.4 ± 0.1 and 2.0 ± 0.1 for how much more likely men are to ask questions than women when there was a female or male chair respectively.

We use Bayes' theorem to assess the statistical significance of these data by comparing a single-parameter model, in which the gender of the speaker or the gender of the chair has no impact on the probability p_F that questions will be asked by women, with a two-parameter model where there is a gender-based effect. We find strong support (a Bayesian evidence or "odds" ratio $R = 11.2$) for a model in which the gender of the speaker has no impact, but we cannot tell from these data if the gender of the chair has any impact (we find $R = 3.2$ weakly favouring the "no difference" model). Figure 3 shows the posterior probability distributions.

This conclusion differs between our data and those collected by Davenport *et al.*, so we ran the same analysis on the data from AAS, as shown in figure 4. Again we find good evidence that the gender of the speaker has little impact on the
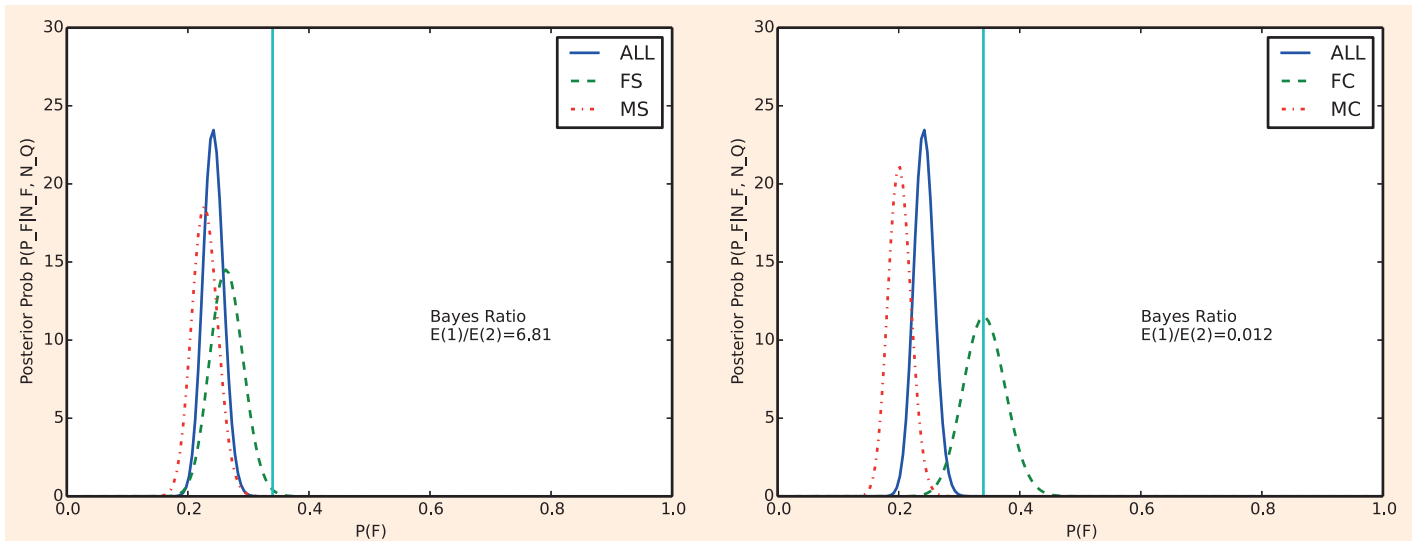
4: AAS 223 data. Posterior probability of p_F, the probability that any individual questioner will be female. Curves are compared for the full conference and for a male or female speaker (left) or chair (right). The vertical line indicates the fraction of female attendees at AAS 223.
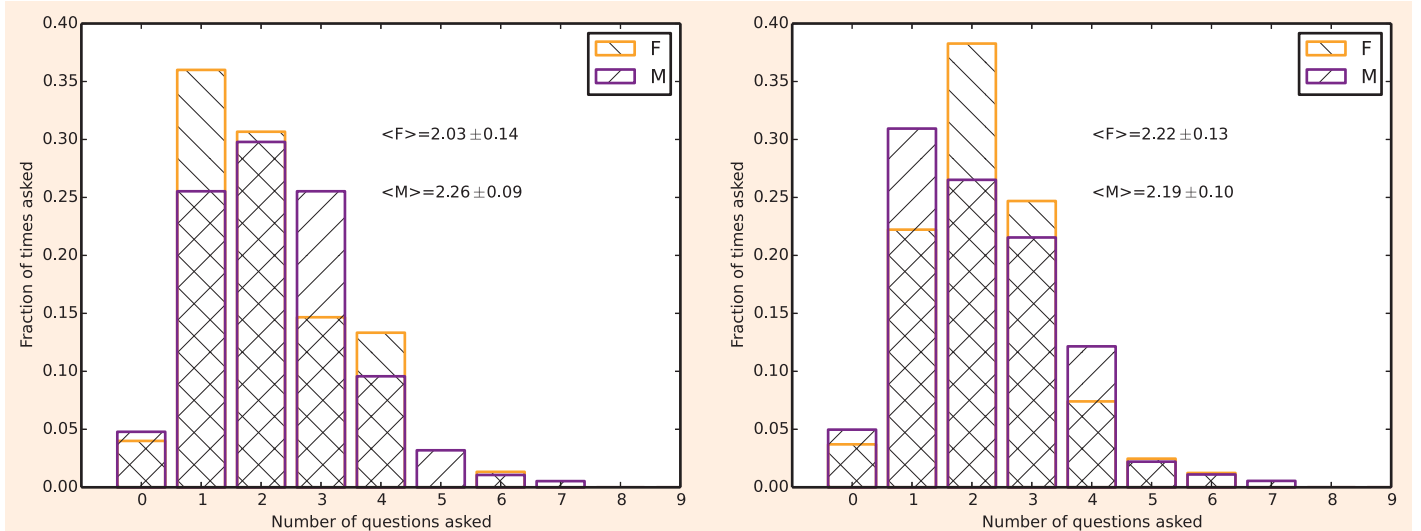
5: NAM 2014 data. Distribution of number of questions asked broken down by gender of speaker (left), chair (right). In each case, the distribution is normalized by the total number of questions at the conference from sessions with a speaker/chair of that gender.

gender balance of the question askers ($R = 7$), but now we favour a difference in the probability of women asking questions when the chair is female versus male ($R = 1/0.012 = 83$ in favour of a difference). In fact, female chairs at the AAS were able to solicit questions from women and men with equal probability (i.e. $p_F \sim f_F$).

On average we find that the number of questions asked at any given NAM talk was 2.2 ± 0.1. We find no detectable difference in the number of questions asked of female and male speakers (both 2.2 ± 0.1), but that female chairs solicited slightly smaller numbers of questions than male chairs (2.0 ± 0.1 for female chairs, compared to 2.3 ± 0.1 for male chairs). The distributions of are shown in figure 5.

Our result on this differs from the AAS survey, which found that female speakers were asked noticeably more questions (3.3 ± 0.2) than their male counterparts (2.6 ± 0.1). We note that these data also reveal that, on average, 0.7 ± 0.2 more

questions are asked per AAS talk than per NAM talk (2.9 ± 0.1 versus 2.2 ± 0.1). This difference might be caused by the different meeting formats (i.e. that talk times are very short at AAS – typically five minutes per talk – and tend to be longer at NAM), or it could reveal a cultural difference between US and UK astronomers.

The survey also allows us to test if men and women typically ask questions at the same time in a string of questions: we find that they do not. The NAM data show that female astronomers were more likely to ask questions later. The median question position for female question askers is 2.26 ± 0.13, while for men it is 1.89 ± 0.05. This difference is largely driven by men being much more likely to ask the first question after a talk.

We show in figure 6 that male astronomers at NAM asked the first question 86% of the time (six times more often than women do), while by the time four questions or more are

asked, women and men are equally represented as question askers – 32% of the 4th/5th/6th/7th questions (aggregated because of low numbers) are asked by women.

Comparison with AAS results

The UK and US astronomical communities have many similarities, but are not the same. As George Bernard Shaw, and later Winston Churchill, said, we are "two nations divided by a common language". In the most recent RAS demographic survey (McWhinnie 2011), the UK astronomy community comprised a total of almost 1700 people, with the proportion of women varying from almost 30% among post-docs to just 7% of full professors in astronomy. The US astronomical community is somewhat larger – the most recent demographic survey of the AAS (Anderson and Ivie 2013) lists 2523 members of whom 25% are female. Internationally, the UK and US astronomical communities

6: NAM data. Gender fraction of questions asked in the *n*th position following a talk. Error bars indicate the central 68% confidence range based on inferring the gender ratio given the observed number of talks. The blue horizontal line at 0.18 indicates the gender ratio of all questions together, while the green line is at 0.28, indicating the gender ratio of all attendees at NAM.
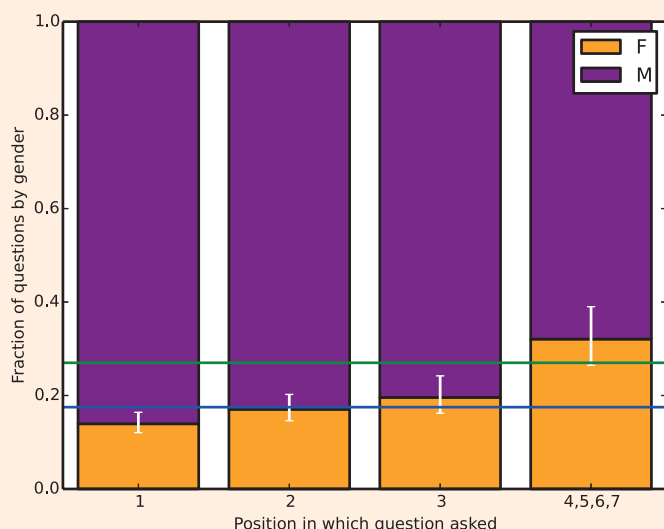
7: The gender balance of NAM attendees, speakers, chairs and question askers compared to the gender balance by professional grade for the UK astronomy community.
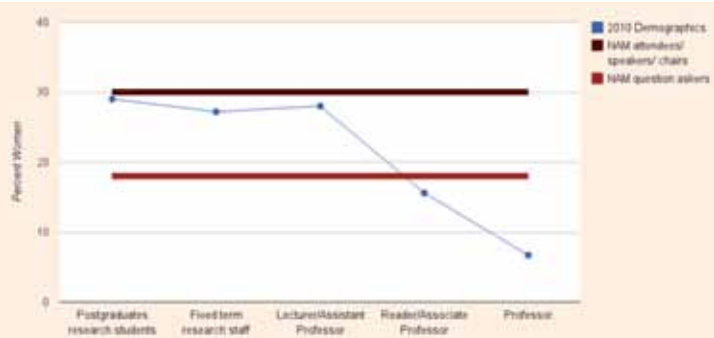
(as tracked by IAU membership; Cesarksy and Walker 2010) are seen as having very similar gender balance, with 11.6% and 12.1% IAU members being women respectively (this relatively low fraction of women among IAU members is usually attributed to the relative seniority of IAU member astronomers).

We find subtle differences in question-asking behaviour at the national conferences of the US (AAS 223) and the UK (NAM 2014). The US community asks slightly more questions on average (0.7 ± 0.2 more questions per talk) and reveals a tendency to ask more questions of female speakers than male, which is not observed in the UK community. In both communities we find that women are less likely than men to ask questions, and that when a session is chaired by a woman this can be improved (slightly as in the UK community, quite substantially as seen in the US data).

The AAS meeting is very different to NAM: it is much larger, and has many more parallel sessions in general. Contributed talks are limited to five minutes and sorted into groups by topic after the abstracts are submitted. The chairs are assigned to parallel sessions by the main organizers, and are required to attend training at the beginning of the meeting, which includes advice on stimulating and moderating discussion following talks. At NAM, by contrast, parallel session topics are proposed to the Scientific Organizing Committee in advance of

abstract submission; abstracts are then submitted to a specific topic session. Once a session is accepted, the proposers (who typically, but not always, act as chair of their session) determine the length of talks, which can therefore vary substantially from session to session. At NAM 2014, the Local Organising Committee emailed session chairs with extensive informal advice on promoting inclusive discussion. All of these differences can change the culture and environment of a meeting in subtle ways, on top of cultural differences between the US and UK.

Implications for UK astronomy

The situation for female astronomers in the United Kingdom has come a long way since 1835 when the Royal Astronomical Society admitted its first women as Honorary Fellows (Caroline Herschel and Mary Somerville; full fellowship was open to women from 1915). Instances of overt discrimination are now thankfully rare in our community (and have been illegal since 1975), but subtle issues such as unconscious bias (demonstrated to affect men and women equally, e.g. Steinpreis *et al.* 1999), as well as stereotype threat (Betz *et al.* 2014, Spencer *et al.* 1999) and sheer numbers can still make the astronomical community feel like a hostile place for women, despite the best intentions of all involved.

It is a fact that less than a third of professional astronomers in the UK are women (McWhinnie

2011); this drops to fewer than one in ten among the most senior (full) professors. There is a rule-of-thumb, that members of a "minority group" will stop noticing they are in a minority when more than about a third of people they interact with is made up of people from the group they identify with. The UK astronomical community may be about to reach this tipping point (and the US community is just past it), so it will be fascinating to see if any changes occur.

We have found that at the NAM 2014, women were around 1.8 times less likely than men to ask questions. A similar observation at the AAS was interpreted as a consequence of question askers being more senior than speakers and attendees in general at the conference (and the lower fraction of women among more senior astronomers). We show in figure 7 that the gender balance of NAM attendees, speakers and chairs roughly matches that found in the UK astronomy community among postgraduate research students, postdocs and most junior academic staff (i.e. lecturers or assistant professors, in the US terminology) as reported in the 2010 RAS Demographic Survey (McWhinnie 2011), while the gender balance of those asking questions at NAM matches the gender balance found among UK-based readers (associate professors).

As a community we should be seeking to make this event a welcoming, diverse and equal opportunity for intellectual discourse about our field regardless of gender or academic status. Our results suggest that either women, and/or more junior people attending NAM, appear to feel less able to ask questions than their more senior and/or male colleagues, which we do not believe is desirable.

The psychology of asking questions

Studies of the psychology of asking questions tend to focus on student participation in classroom discussions (e.g. Krupnick 1985, Younger *et al.* 1999, or see Murphy and Whitelegg 2006 for physics-specific classrooms), or contributions to conversation in general (e.g. see the review by James and Drakich 1993) rather than question-asking at professional research conferences. Nevertheless, it appears that a lower engagement from women than men is a fairly ubiquitous finding in such studies, which tend to conclude that men/boys on average dominate most kinds of mixed discussions (James and Drakich 1993, Krupnick 1985, Younger *et al.* 1999), and if women participate equally they risk being perceived (negatively) as dominating the conversation. We consider here if any general conclusions or ideas for good practice can be drawn from this rich social science literature on gender and discourse.

These sources include much discussion of creating an environment conducive to participation so that people feel "psychologically safe in taking the risk of asking a question". Research

on the impact of stereotype threat, or the risk of being judged by negative stereotypes of the ability of women in maths and science (e.g. Spencer *et al.* 1999) do support the idea that female astronomers (on average) would tend to find it harder to ask questions. For example, Carr and Steele (2010) found that women facing stereotype threat are more risk averse than typical. In addition, there are suggestions that subtle differences in criticism from teachers over time can lead girls/women to develop low self-confidence in their abilities, while boys/men may (again on average) overestimate theirs (e.g. Dweck 1986). This boost in self-confidence given to male astronomers over the course of their schooling may make them more likely to be willing to ask questions. Female astronomers (on average) may simply need more encouragement and help to ask questions.

Sunderland (2000) cautions against the representation of women as "victims" of male dominance in discourse, for example pointing out there is a difference between quantity and quality of interactions, and reporting in her work (based on observations of language learning classrooms) that it was a just a small fraction of the boys who were dominating the discourse. Judging the quality of questions following a talk would be difficult and risks subjective assessment (prone to unconscious bias). We suggest the length of response from the speaker could be taken as an imperfect but objective measure. It would also be difficult to record who is asking the questions, unless questioners were encouraged to identify themselves by name at the start of their question.

Studies have found that the presence of a female role model has a positive impact on female participation (e.g. Krupnick 1985). In our NAM survey, a female chair had only a mild impact on the gender balance of questioners; the AAS survey found that a very significant difference. Based on these results, a good gender balance of session chairs can only improve the environment that encourages women to ask questions. Favouring women for the first question may also have a simple positive impact.

In classroom interactions, positive effects on gender equal participation have been shown when teachers wait longer to get answers and don't immediately go to the first person who raises their hand (Murphy and Whitelegg 2006). It's curious that in the NAM data we show women are as likely as men to ask later questions. Providing session chairs with strategies to maintain discussion and encourage longer Q&A sessions may have a positive impact on the participation of female astronomers in asking questions.

Suggestions for action

The findings of this first survey of the gender balance of questioning at NAM may have raised more questions than it answered. While we detect a clear gender difference in attendees and those asking questions, given what we know about gender balance changes with age/seniority in UK astronomy, we are unable to tell if this is caused directly by gender differences or by age/seniority differences in question-asking behaviour. If this survey is replicated in future it would help to have access to both the gender balance and the seniority of attendees of NAM. We found that women were more likely to ask later questions in a string, but did not record how often Q&A sessions were cut short before all questions were asked and if this behaviour had any impact on the overall gender balance. We also did not record if the chairs routinely ask questions, and if this behaviour shows any gender differences.

We propose the following set of actions based on the findings of this work and our review of the literature on the psychology of question asking:

(1) A similar survey should be repeated at future NAMs (perhaps not every year, but at least within the next few years), and monitored for changes. In that way we can test if our suggested actions have any impact, as well as disentangle the roles of gender and seniority in asking questions. In future surveys we recommend that the following additional data be collected:

● A unique identifier for each talk (preferably assigned beforehand to enable a drop-down menu in the form).

● If the chair asked a question, and when (i.e. collect a string such as "FFCM" where C represents a question asked by the chair).

● Was the Q&A session cut short, or did it end due to lack of further questions?

● If possible, record the name of the question asker as well as their gender.

● The gender and seniority of NAM attendees.

(2) Chairs of sessions should be given a brief training session or sent guidelines with advice for good practice, which we suggest should include the following recommendations:

● Younger scientists should be explicitly encouraged to ask questions (i.e. this should be stated in introductory remarks by the chair), and favoured if there is a choice of questioners.

● If there is a choice between male and female questioners for the first question, a question from a woman should be given priority.

● Questioners should be asked to identify themselves by name.

● If possible, Q&A sessions should not be cut short before at least four questions have been asked (if they need to be ended early). To enable this, session organizers should schedule enough time for questions and speakers should not be allowed to run over time. We believe these actions will help to make our annual meeting a more open opportunity for discourse among professional astronomers regardless of their gender or seniority. ●

Jonathan Pritchard, Imperial Centre for Inference and Cosmology, Imperial College, London, UK; j.pritchard@imperial.ac.uk. Karen Masters, Institute for Cosmology and Gravitation, University of Portsmouth, UK; karen.masters@port.ac.uk. James Allen, Sydney Institute for Astronomy (SIfA), University of Sydney, Australia; j.allen@physics.usyd.edu. au. Filippo Contenta, Department of Physics, University of Surrey, UK; f.contenta@surrey. ac.uk. Leo Huckvale, Jodrell Bank Centre for Astrophysics, University of Manchester, UK; leo. huckvale@postgrad.manchester.ac.uk. Stephen Wilkins, Astronomy Centre, University of Sussex, UK; s.wilkins@sussex.ac.uk. Alice Zocchi, Department of Physics, University of Surrey, UK; a.zocchi@surrey.ac.uk.

References

Anderson G and Ivie R 2013 Demographics survey of 2013 US AAS members, summary results https://aas.org/files/resources/aas_members_workforce_survey_final_jan2014v2.pdf.
Betz D E et al. 2014 *The SAGE Handbook of Gender and Psychology, Chapter 26.*
Bowler S 2014 *Astron. and Geophys.* **55**(4) 10.
Carr P B and Steele C M 2010 *Psychological Science* **21**(10) 1411.
Cesarksy C and Walker H 2010 *Astron. & Geophys.* **51**(2) 33.
Davenport J R A et al. 2014 *Studying gender in conference talks – data from the 223rd meeting of the American Astronomical Society* arXiv:1403.3091.
Dweck C S 1986 *American Psychologist* **41**(10) 1040.
James D and Drakich J 1993 Understanding gender differences in amount of talk: a critical review of research, in *Gender and Conversational Interaction* ed. D Tannen (Oxford University Press) 281.
Krupnick C G 1985 *On Teaching and Learning* **1**(1) 18.
McWhinnie S 2011 *The Demographics and Research Interests of the UK Astronomy and Geophysics Communities 2010* http://www.ras.org.uk/images/stories/ras_pdfs/Demographic_Survey_2010_-_final_report.pdf.
Murphy P and Whitelegg E 2006 *Girls in the physics classroom: a review of the research on the participation of girls in physics* (Institute of Physics, London).
Simpson R 2014 *Astron. and Geophys.* **55**(4) 15.
Spencer S J et al. 1999 *Journal of Experimental Social Psychology* **35** 4.
Steinpreis R et al. 1999 *Sex Roles* **41** 509.
Sunderland J 2000 *Language Teaching Research* **4**(2) 149.
Younger M et al. 1999 *British Journal of Sociology of Education* **20**(3) 325.