

Sport and the leaky pipeline

Published 24 days ago by [Joe Smith](#)

"It is important that we know where we come from, because if you do not know where you come from, then you don't know where you are, and if you don't know where you are, you don't know where you're going. And if you don't know where you're going, you're probably going wrong."^[1]

J.K Penhau Smith

Introduction

This month we are looking in more detail about the kind of things we can start to understand when we have "good" data. This is a variation on the kind of work which [UHIWWC completed by interviewing international students](#) and hinges on a background [understanding of the population of people from which our sports club is drawn](#). This week the focus is on the "leaky pipeline" and how we might measure its existence.

When we look at a population of individuals in a given area (geographic, intra-club, inter-club, within a sport etc), arranged into a hierarchy of achievement, we can create a mental model, a "pipeline" of progress. The pipeline metaphor was originally created to be applied to academia and has been focussed on getting people into higher levels of Science, Technology, Engineering and Maths (STEM), from undergraduate through to tenured professor. This diagram is classically drawn as a pyramid, with people "leaking" out of the pipeline as they leave academia (Fig: 1). These "leaks" are not intrinsically good or bad. Not everyone wants to be a senior academic and society would be very different if they did. Similarly, people might get to a level and stay in that pool of people, rather than leaving the system. For example, a lifetime career as a post-doc is very possible and might be the aim of an individual.

Diagram Description automatically generated

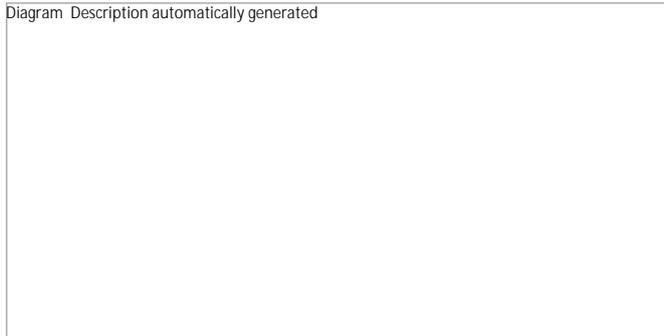


Fig 1: the "leaky pipeline" of academia

The reason that this model is interesting is that it can be used to measure the "leaks" in the system and, when combined with an understanding of the challenges faced by people moving up the pyramid, can be used to visually show how these barriers affect who actually rises to the top of the pyramid.

These "leaks" can be seen if you have good data about the population. All we need is the to know the population we are drawing from (in this case everyone who goes to university) and some basic information about the factors which we are interested in (classically gender or ethnicity). As a result, we can definitively say that women tend to leak out of the pipeline at a greater rate than men in STEM (Cronin and Roger 1999; Resmini 2016). This has then led to research into why this might be the case, the "barriers" faced by those rising from one level to another and in theory it is possible to then address these barriers.

So, does this also occur in sport? And can we show this using what we have learnt about producing "good" data?

In short yes, there is a "leaky pipeline" of sport participation, from beginner to Olympic level, across a range of sports (Curran et al. 2019; Zarrett et al. 2020). If this happens across a sport as whole, can we also observe it a sports club?

Let us conduct a purely hypothetical thought experiment to find out. There will be some maths in this blog, but I am pulling numbers out of thin air to make the maths easy, so this data should not, under any circumstance, be considered to be representative.

Let us imagine that we have a sailing club of 1000 people. We are then going to ask them at what level they have most recently participated and also what their favourite colour is: magenta, cyan or yellow. We can then plot these data on a figure which accounts for the proportion of our population. The data I made up to make these figures is here: and everything was put together in Microsoft Excel (2010), although other programs are available.

As we expect, the majority of participants identify as beginners, while the number of participants at each level declines, finally down to one person going to the Olympics from this club (Fig: 2). When we ask about their favourite colour, in absolute terms there are a few people more overall who preferred the colour magenta, compared to cyan or yellow (365 compared to 318 and 317 respectively). And this might only translate to a few more people who prefer magenta in absolute terms across participation level (Fig: 3). When we look at the colour preference across participation level, accounting for the changes in the number of people at each level, by putting everything in terms of the percentage of participants, we see a consistent domination of people who prefer the colour magenta (Fig: 4).

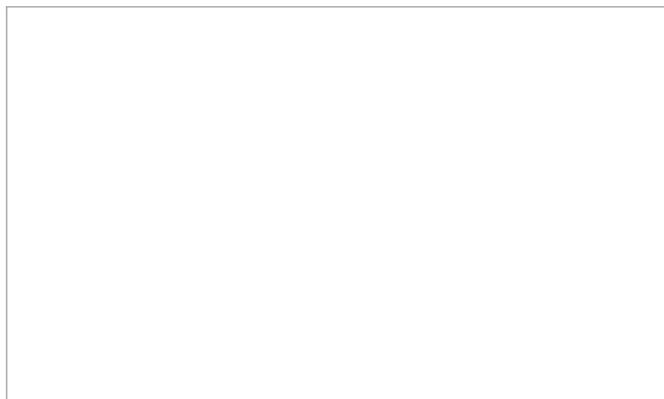


Fig 2. Total number of participants, by level of participation at our sailing club.

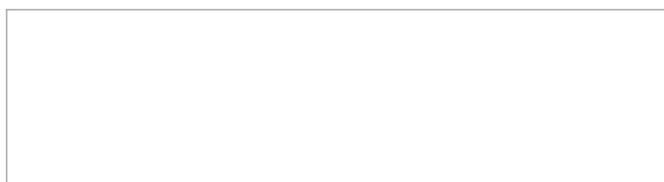




Fig 3. Number of people who prefer the identified colours, by participation level.



Fig 4. Proportion of people who prefer a given colour, by participation level.

Is this an issue? We can see that our starting population of beginners has no preference for colour, so if there were no factors which affected people's move up the participation level, then we would expect that the ratio of magenta: cyan: yellow preference would not change beyond the measurement error (something we will talk about in a future blog).

This is not the case in this example, with the increase in the proportion of people who prefer the colour magenta across the higher participation levels. This suggests that there might be an unmeasured factor(s) which is preferentially selecting from those who prefer magenta, causing them to rise through the participation level. Alternatively, there may be barrier(s) to participation only apparent to those who prefer the other colours. Since there is no "prior" reason to think that liking one colour over another might make someone a better sailor, where sailing is concerned^[2] then we can suggest that when moving up the participation ladder, colour preference is implicitly or explicitly being selected for, or barriers are being placed in the way of those who prefer other colours.

Does this matter? This is only a point of data, so a longer-term survey pattern would be critical to be sure that we are seeing an over-representation of those who prefer magenta at the higher levels of participation. How we might deal with this information and how we can account for our measurement errors are subjects which will be dealt with in later blogs.

If this corresponds to other pieces of evidence that also observe an over-representation of colour preference in this sport, then we can consider that this might be an issue, as we are not seeing sailors with other colour preferences being represented at high levels, meaning we cannot gain their insights or contributions to the sport or the club. Therefore, we now have a baseline from which we can say that our interventions to help increase representation of those with other colour preferences are working or not, and how effective they may be.

So, what does this mean when we were to look at a real-world example? While colour preference may not be that relevant for the majority of sports clubs, we can use this method as a way to look at how other factors, such as gender or ethnicity might influence progress up the participation levels in a sports club. This is a similar method to the one used by UHIWWC to understand at what level their international students were participating at. They followed up by interviewing those sailors who were not making it to higher participation levels, as these were the ones facing barriers. By doing this they understood the barriers faced and could take action to address these issues, improving their participation at higher levels, meaning they outperformed teams of equivalent experience levels. Not only that, but they could absolutely say that their intervention had worked to improve their participation and they could acquire more funding on this basis from external funders.

Conclusions

This month we looked at how we might start to use the information that we are able to gather to start to draw conclusions about our sports clubs. We looked at the "leaky pipeline" model and how we can arrange our club into a hierarchy of participation. We can also subdivide these groups based on other factors which we can measure, using those principles of "good" data that we have previously seen. As a result, we can visually represent over-representation of sub-populations at given participation levels and understand areas of focus for us to improve the diversity of our sport.

References

These papers are mostly open access, but if you would like more information please do contact me.

Crawley S (1998) Gender, class and the construction of masculinity in professional sailing: A case study of the America 3 Women's Team. *Int Rev Social Sport* 33:33–42

Cronin C, Roger A (1999) Theorizing progress: Women in science, engineering, and technology in higher education. *J Res Sci Teach* 36:637–661. [https://doi.org/10.1002/\(SICI\)1098-2736\(199908\)36:6<637::AID-TEA4>3.0.CO;2-9](https://doi.org/10.1002/(SICI)1098-2736(199908)36:6<637::AID-TEA4>3.0.CO;2-9)

Curran O, MacNamara A, Passmore D (2019) What About the Girls? Exploring the Gender Data Gap in Talent Development. *Front Sport Act Living* 1:10. <https://doi.org/10.3389/fspor.2019.00003>

Low V, Dillion L, Caffari D, et al (2019) Women in Sailing Strategic Review. London

Resmini M (2016) The 'Leaky Pipeline'. *Chem - A Eur J* 22:3533–3534. <https://doi.org/10.1002/chem.201600292>

Zarrett N, Veliz P, Sabo D (2020) Keeping Girls in the Game: Factors that Influence Sport Participation. New York

If you enjoyed this you can find all my other ConnectedCoaches blogs - including the others in this data series - [here](#).

Login to follow, share, comment and participate. Not a member? [Join for free now.](#)

[1] Pratchett. T, 2010, *I Shall Wear Midnight*, edition: 2010, Random House publishing, London

[2] For more information on factors which might influence participation in sailing please see Crawley 1998 and Low et al. 2019 amongst others.

tags : planning, learning, development