Approach Overview

This document provides a summary of the research approaches we are planning to take in the empirical strands of the Action Research for Research Culture project. It will develop and evolve as the project does.

1. Narrative CVs

Core experiment

We will test the hypothesis that the Résumé for Research and Innovation CV format alters the assessment of candidates at the shortlisting for interview stage of postdoctoral recruitment.

We are considering two randomised controlled designs – see Option A and B below – and currently have a strong preference for the first design.

Option A: A parallel randomised self-controlled trial

Candidates will be asked to prepare a narrative CV in addition to their standard CV. This will require extra work on the part of the applicants but we hope it will also provide benefits to them. We will provide personalised feedback on their application from a careers adviser; they are likely to have to prepare such CVs in future, so will be able to reuse what they write; and this application process will provide them with a safe environment for their first attempt, as their standard CV will be assessed alongside their narrative CV in the final selection.

As shown in Figure 1 below, selection panels will be divided in two (orange and red), with one half being given the narrative format CV (green) and one half being given candidates standard CV (blue) (stage 1). Once the initial ranking has been completed, all panel members will be provided with both forms of CV and asked to revisit their rankings and complete the shortlisting process (stage 2).

![Figure 1: Experimental design for self-controlled narrative CV experiment](image)

Option B: A cluster randomised controlled trial. This method will randomise the CV format in each recruitment process. For half of the recruitments we will request standard CVs, for the other half narrative CVs. Recruiters will rank applicants as before. This method is logistically easier, doesn’t require two ranking stages and only requires one format of CV from each applicant; however, it will be less sensitive to differences (as there will be more noise because the CVs of different candidates are being compared) and it provides no ‘safety net’ for either candidates or recruiters using narrative CVs for the first time. Importantly it also it requires the development of scales to rank candidates on a range of competency criteria (to provide a comparator for the performance of the different CV types) -
this step is required only for the secondary analysis of Option A. For these reasons we prefer, and will initially trial, Option A.

Carrying out an experimental comparison of standard and narrative CVs allows us to contrast them both quantitatively (by analysing ranking data) and qualitatively by analysing the content of the CVs and interviewing/surveying the academics who are recruiting and the applicants who are applying. All these data sources are likely to be equally important.

**Primary outcome**

We will fit a latent variable model to see if narrative CVs lead to the selection of different applicant characteristics (modelled as a single predictor), to those characteristics that are selected when standard CVs are reviewed (Figure 2).

![Figure 2: Simplified variable model (latent variables circled)](image)

We measure ranking (R) for standard and narrative CVs and assume they are driven by two latent variables: one measure of narrow academic achievement (A\text{aAdv}) and one measure of wider contribution (A\text{wAdv}). We normalise the influence of academic achievement and then use the rankings to estimate the difference between the two latent variables, \( \alpha \) and \( \beta \). We will then ask if the type of difference detected is consistent with the intent of narrative CVs: that of improving research culture and increasing the diversity of researchers shortlisted. We will test this by looking at whether the rankings of underrepresented groups (female applicants, minority applicants) increase with narrative CVs.

We will also test whether any ranking change is affected by potential confounders such as prior knowledge of the candidate or the ranking position – is it that lower rankings are changed by narrative CVs, but the top rankings aren’t.

To test whether narrative CVs affect the confidence of shortlisting we will ask a subsample of recruiters to provide rankings of candidates and score their confidence in each of the rankings.

**Qualitative exploration**

In addition to the ranking data, we will collect qualitative feedback through a mixture of semi-structured and ‘Think aloud’ interviews from a sample of applicants and recruiters to understand preparation/assessment of the applications and what other sources of information they refer to. These interviews will allow us to explore how candidates and panel members experience alternative formats in terms of ease of use, time invested, etc. and what tools, templates and training would facilitate adoption of the new formats. We will select interviewees initially to ensure the maximum diversity of our sample and then use the initial results and quantitative work to purposively select additional interviewees to ensure we minimise the number of interviews necessary.

**Secondary analysis**

To understand what is driving the differences identified in the primary analysis, we will develop a series of ‘competency criteria’ and ask recruiters to score the candidates against these criteria. Criteria for the assessment will be developed from a hybrid of Vitae’s Researcher Development Framework,
Yuki’s 2012 hierarchical framework of development behaviours and the Glasgow Narrative CV pilot. We will then statistically analyse these scores to identify whether they are components of the ranking scores and/or the latent variables. Alongside this we will look to see if we can predict the rankings or the academic achievement latent variable from bibliometric or administrative information (Andersson et al., 2021; Holst and Hägg, 2018). This analysis, alongside the qualitative exploration, will help us to understand which key competencies drive the evaluation of narrative CVs. Carrying out this analysis as a secondary aim will allow us to start the initial experiments before we complete the scale development for the competencies and will also allow us to test whether there is a separate effect candidate assessment caused by asking the recruiters to score on structured scales, as some previous research indicates that this can reduced biases (2012).

We will also be able to analyse how shortlisting ranking affects the likelihood of recruitment to estimate what size of changes are needed in rankings to affect final recruitment decisions.

Sample sizes
The paucity of previous experimental recruitment work means we can find no directly comparable studies - this makes up-front estimates of likely effect sizes and/or power calculations impossible. Norman et al (2012) recommend that in such circumstances heuristic estimates of sample size should be used.

Analysing the University of Cambridge recruitment system from 2017–2019 shows an average of 470 postdoctoral recruitments per year which attract an average of 20 applicants. If we estimate being able to enrol 3-5% of recruitments into our study over the two and a half years (starting lower and ending higher) this would give us a sample size of between 700 and 1200 applications which heuristic estimates suggest is likely to be sufficient to detect meaningful differences in shortlisting. These numbers will allow us to analyse whether any effects are particularly pronounced for, or differentially affect, under-represented research groups.

Disciplinary differences
Résumé for Researchers is being suggested by UKRI as a common format across all disciplines, so we think it is important to test it across the full variety of disciplines. This will give us valuable qualitative information on using the format even in areas with relatively small numbers of recruitments. To address this problem quantitatively we will take two approaches:

- Fit a multi-level model to allow the effects of narrative CVs to vary by discipline
- Fit single level models in disciplines where we have the largest number of recruitments enrolled.

Although the number of recruitments is substantially lower in the Arts and Humanities – the average number of applicants is relatively constant across all areas except the Physical Sciences where it is higher (Figure 3).
2. Professional Development Tools

This strand will identify the strengths and weaknesses of a series of tools and approaches that facilitate development/feedback conversations, and their applicability to different researcher development needs. To allow a comparison of the tools we will incorporate pairs of tools into professional development courses and activities in our institutions, see Figure 4 for a graphical overview. Researchers embarking on professional development training will be surveyed to identify the areas they are keen to develop. As part of their training we will include two alternative development tools. At the conclusion of the training we will capture Likert ratings for how useful they found each tool for each development need alongside qualitative reasons.

![Figure 4: Outline method for development tool comparison](image)

We will base these Likert scales surveys on semi-structured interviews with initial course participants to understand their development needs and elicit their views of two alternative tools. These survey tools will then allow us to capture similar information from a larger number of participants as the project develops.

Previous work carried out with a cohort of UKRI Future Leader Fellows and their coaches identified the following recurrent themes for development that we will use as an initial structure to classify researcher development needs:

**Team dynamics:** Performance management; Delegation; Managing dispersed teams; Conflict resolution; Communication and collaborative relationships; Preventing bullying / harassment

**Personal effectiveness:** Wellbeing; Confidence / vulnerability / fear of failure / assertiveness; Coaching skills; Email / diary / time management; Influencing without authority

These themes will be compared with those identified in our international and industry partners to identify additional potential themes to be explored.

The tools we will test include, for example: use of a formal 360 (as developed and currently in use with the Future Leader Fellows development network); the Vitae RDF planner; the ‘GROW’ coaching model; the ‘SUBIC’ feedback model; careers toolkits provided by professional bodies, e.g. the Royal Society of Chemistry; MBTI profiling; and institutional annual appraisal paperwork.

The tools will initially be trialled for two groups: postdocs, and early-career Principal Investigators (PI) (new lecturers, career development fellows). Alongside information on the individual’s experience of the tools we will collect baseline data to understand how variations in local settings might affect an individual’s experience, e.g. a department with a strong existing culture for appraisal. We will use this information to produce guides for researchers on which tools are most appropriate for different development needs and situations, alongside a set of hints and tips for applying the tools in an academic setting.
3. Precarity, redeployment and contract extensions
This strand aims to empirically understand the extent to which redeployment could provide a solution to early career researcher precarity and whether contract extensions might increase the feasibility of redeployment by providing a longer time window or looser skills match requirement for transfer between contracts. It will also provide quantitative and qualitative data to test other approaches to addressing the issues caused by precarity of researcher employment.

We will investigate the potential of precarity through four approaches:

Determining what researchers consider attractive ‘redemption’
We will conduct semi-structured interviews with postdocs at different stages of their contracts to understand how they choose career options, to determine how redeployment could contribute and what they consider the boundaries between ‘redemption’ and ‘career change’. We will initially interview postdocs from a range of disciplines to determine the level of common ground in their career aspirations and whether we need to develop discipline specific instruments for subsequent stages of the research.

We will use these interviews to design a stated-preference discrete-choice experiment to determine the relative importance attached to different characteristics (duration of appointment, subject, location, time remaining on current contract etc) when seeking a new role. The discrete-choice experiment will be conducted on-line (with a tool such as conjoint.ly) with postdocs recruited from partner institutions and through the Researchers 14 network – allowing us to include both a broad UK and international sample of researchers. Cambridge has a population of over 4,000 post-docs and there over 50,000 post-docs in the Researchers 14 network which will allow us to achieve a sample size that is large enough to investigate the different characteristics of different disciplines and the impact of different personal situations/characteristics (such as relationship status and caring responsibilities).

The discrete-choice approach will provide us with insights into how postdocs trade off alternative appointments/redeployment opportunities. Discrete choice methods have been used successfully to investigate post-doc career choice – but only examining a more limited set of factors (institutional characteristics, gender and ethnicity) but not used to examine subject preference, duration and location (Pinheiro et al., 2017).

Investigating how PIs select applicants and the potential for training to allow increased redeployment
We will explore how recruiting academics’ select candidates and their views of redeployment by interviewing a sample of PIs/managers to understand the key drivers of selection and we may carry out a similar state preference approach to quantify these preferences. We will explore whether, and how long, PIs would be willing to postpone appointment to allow internal candidates time to be trained, either externally or on-the-job.

Estimating level of vacancies suitable for redeployment and the impact of contract extensions
We will analyse the University of Cambridge’s historical vacancy information to estimate the likelihood of suitable vacancies for redeployment arising within different time windows for researchers in different fields and with different personal circumstances, using the data from the state-preference experiment to estimate the attractiveness of other posts. This will allow us to estimate the acceptability and feasibility of redeployment at end of contract and the impact of contract extensions on increasing the possibilities for redeployment.

To provide empirical data on extensions we will investigate the current use of contract extensions supported by departments and PIs. We would use HR data to identify instances of contract extension and carry out semi-structured interviews with PIs and Postdoc to explore how these extensions came
about and what benefits they provided. This approach suffers from more selection bias than an experimental approach but will provide valuable insights into the length of extensions that might be necessary to see effects and the possible outcomes.

4. References


Norman, G., Monteiro, S., Salama, S., 2012. Sample size calculations: should the emperor’s clothes be off the peg or made to measure? BMJ 345, e5278–e5278. https://doi.org/10.1136/bmj.e5278
