

technopolis_{|group|}

July 2016

Evaluation of the ESRC Transformative Research Scheme

Final Report

Evaluation of the ESRC Transformative Research Scheme

Final Report

technopolis **group** July 2016

Peter Kolarz

Erik Arnold

Kristine Farla

Stephanie Gardham

Kalle Nielsen

Cristina Rosemberg-Montes

Martin Wain

Table of Contents

Executive Summary	1
The Scheme in context	1
Transformative research	1
Definition.....	1
Success.....	2
Profile and Attractiveness of the Scheme	2
Function and operation of the assessment stages.....	3
Anonymous Review (Sift Panel)	3
Presentation Event (Pitch-to-peers).....	3
Establishing a culture of transformative research	3
Outputs, outcomes and future perspectives	4
Summary of Recommendations.....	5
Introduction	6
1 The Scheme: description, context and comparison	8
1.1 The ESRC scheme	8
1.2 What is ‘transformative research’?.....	10
1.2.1 Two notions of transformative research: transforming science or transforming society	11
1.3 Funding transformative research: problems and solutions	12
1.3.1 Risk and conservatism.....	12
1.3.2 Review panels and remote peer review: expertise and strategy	14
1.3.3 Funding people or funding projects.....	15
2 A successful scheme for a slippery concept	17
2.1 The meaning of ‘transformative research’: answers from the Scheme.....	17
2.1.1 Towards an expert definition	18
2.2 Main findings from the project review exercise.....	19
2.2.1 Defining ‘transformative research’	19
2.2.2 The effectiveness of the Scheme	19
2.2.3 A pluralistic process for a pluralistic concept.....	20
3 The assessment stages: individual significance	21
3.1 Stage 1: The institutional level.....	21
3.1.1 Proposals submitted to the scheme: overall characteristics	23
3.2 Stage 2: Sift panel	24
3.2.1 The sift panel of the 2014-15 call	26
3.3 Stage 3: Pitch-to-peers.....	28
3.3.1 The function of Pitch-to-peers	30
4 The three calls compared: understanding the Scheme as a learning process	33
4.1 Ensuring a level playing field	33

4.2	Rising consensus between panellists and fellow applicants at Pitch-to-peers events	34
4.2.1	Feedback on Pitch-to-peers.....	35
4.3	Other learning curves and feedback loops	37
4.3.1	Reflexivity of the assessment panel	37
5	Perception and attractiveness of the Scheme	39
5.1	Individuals.....	39
5.1.1	Lack of wider context and integration	39
5.2	Institutions.....	40
5.2.1	The 2012-13 institutional payments	40
5.3	Success rates.....	41
6	Phase 2: Outputs, Outcomes and future perspectives.....	43
6.1	Productivity of the grants	44
6.1.1	A note on unsuccessful grants.....	46
6.2	The transformative character of the grants	46
6.2.1	Transformative qualities in detail.....	48
6.2.2	Transformative outcomes in context	50
6.3	Determinants of transformative scope.....	51
6.3.1	The importance of quality	51
6.3.2	The seniority-problem.....	52
6.4	Further plans and hindsight-perspectives	53
7	Conclusion and recommendations.....	56
7.1	Taking the Scheme forward.....	56
7.2	Funding Transformative Research: A re-appraisal	56
7.3	List of recommendations	57
7.3.1	Headline recommendation: the Scheme should continue.....	57
7.3.2	Operation	57
7.3.3	Integration	58
7.3.4	Beyond the Scheme	58
	Appendix A Details of methodology	59
	Appendix B Bibliography.....	67
	Appendix C Statistical analysis of Phase 1 applications review data	69
	Appendix D Comparator schemes	77
	Appendix E Observation notes: Sift Panel 22-04-15.....	79
	Appendix F Observation notes: Pitch-to-peers 17/18-07-15	82
	Appendix G Comparing Panellists' and fellow applicants' Pitch-to-peers scores	85
	Appendix H Survey questions	88
	Appendix I Phase 2: Interview details	93
	Appendix J Phase 2: Details of Phase 2 review exercise.....	96

Tables

Table 1: Sift panel – assessment criteria.....	9
Table 2: Details of applications.....	10
Table 3: 2012/13 call submissions – fit with ESRC strategic priorities	11
Table 4: Overview of other transformative research funding schemes.....	12
Table 5: Extreme opposite mind-sets to risk-taking.....	13
Table 6: Reviewers’ score as predictor of outcomes of the Scheme	20
Table 7: Association between sift panel and Pitch-to-peers scores	20
Table 8: Outcomes of institutional selection processes.....	21
Table 9: Profile of questions asked at Pitch-to-peers event.....	31
Table 10: Character of questions – applicants and panellists compared.....	32
Table 11: Success rates of the 2014-15 call by gender and seniority.....	34
Table 12: Comparing Panel and fellow applicant mean scores (out of 10).....	35
Table 13: Disadvantage of presenting late in the day.....	36
Table 14: Scheme success rates and contrast to ESRC totals	41
Table 15: Mean outputs listed in ResearchFish.....	44
Table 16: Review exercise – verdict on transformative character (individual review level)	47
Table 17: Peaks – percentages of reviews scoring the highest possible mark.....	49
Table 18: Correlations between markers of quality and definitions of transformative research.....	52
Table 19: Negative correlations between early career research and quality/ TR scope	53
Table 20: Interviews – successful and transformative results	54
Table 21: Survey – Respondent Details	59
Table 22: List of interviewees.....	60
Table 23: Sample of proposals for the review exercise	62
Table 24: List of reviewers for the project review exercise.....	63
Table 25: Verdict.....	70
Table 26: Predicted probabilities of a positive verdict, given value of scores.....	71
Table 27: Predicted probabilities of a positive verdict, given value of scores	72
Table 28: Regression analysis (Ordinary Least Squares), <u>at project level</u>	73
Table 29: Regression analysis (Probit), <u>at individual level</u>	76
Table 30: Details of other transformative research funding schemes.....	77
Table 31: Sift Panel – Thematic summary of comments and discussion points.....	81
Table 32: Sift panel – Comments on successful proposals.....	81
Table 33: Sift panel – Comments on rejected proposals	81
Table 34: Comparing Panel and peer scores for the 2012/13 call	85
Table 35: Comparing Panel and peer scores for the 2013/14 call	86
Table 36: Comparing Panel and peer scores for the 2014/15 call	87

Table 37: List of Phase 2 interviewees	93
Table 38: Summary of interviews	94
Table 39: Review exercise – mean and median figures across the criteria (TR/ standard split).....	99
Table 40: Overall verdicts	100
Table 41: % of reviews that received the highest possible mark.....	101
Table 42: Differences between reviews for 2012-13 and 2013-14 TR outputs.....	102

Figures

Figure 1: Selection steps and key considerations	8
Figure 2: Transformative research – Challenges and funding solutions	14
Figure 3: Survey - Define 'transformative research'! (n=78).....	17
Figure 4: Survey – Internal selection process (n=74).....	22
Figure 5: Survey – Accessing the Scheme (n=78)	22
Figure 6: Mean scores for all projects: top and bottom three categories.....	23
Figure 7: Survey – Should the practice of anonymous review be expanded to other funding schemes? (n=77)	24
Figure 8: Survey – Chance of success; effects of anonymity (n=77)	25
Figure 9: Survey – Views on two-page format (n=77)	25
Figure 10: Sift panel – Thematic summary of comments and discussion points	26
Figure 11: Sift panel – percentages of proposals that triggered disagreement between reviewers (by theme)	27
Figure 12: Sift panel – positive and critical comments on successful and rejected proposals (by theme).....	28
Figure 13: Survey – Feedback on Pitch-to-peers workshop 1 (n=37).....	29
Figure 14: Survey – Feedback on Pitch-to-peers workshop 2 (n=37)	30
Figure 15: Success rates of the 2012/13 call by gender and seniority.....	33
Figure 16: Success rates of the 2013/14 call by gender and seniority	34
Figure 17: Correlation between Peer and Panel scores at Pitch-to-peers events	35
Figure 18: Survey – Pitch-to-peers Improvement 2012/13 vs. 2013/14 (mean scores out of 5).....	36
Figure 19: Main learning and feedback loops.....	38
Figure 20: Survey – Further comments on the Scheme: Themes and frequency.....	39
Figure 21: Aberystwyth University Transformative Research Symposium	41
Figure 22: Review exercise – quality profiles of TR and Standard grants outputs (individual review level).....	45
Figure 23: Consensus and disagreement at the level of projects (question: Overall, would you describe this output as 'transformative')	48
Figure 24: Review exercise – areas of greatest difference between TR and Standard grants (individual review level)	49
Figure 25: Mean level of disagreement between reviewers – most and least contentious statements.....	63
Figure 26: Distribution of Score Pitch-to-peers panel per final verdict, <u>at project level</u>	74
Figure 27: Percentage of reviewers that gave a positive final verdict, <u>at individual level</u>	75

Executive Summary

ESRC commissioned Technopolis to evaluate its Transformative Research Scheme, which has so far run three calls (2012-13, 2013-14, 2014-15) and awarded 45 grants of around £200,000-£250,000 each for research projects lasting 18 months.

The Scheme has two unusual elements within its assessment process for applications: first, an anonymous panel review stage, i.e. panel reviewers have no information about the applicant and are limited to a two-page description of the proposed project itself; second, a presentation event ('Pitch-to-peers'), where shortlisted applicants present and discuss their proposal with assessment panellists and fellow applicants.

Aims of the evaluation are to assess whether the Scheme successfully identifies and funds research that is transformative, analyse operational aspects of the Scheme and inform future development and wider funding activity, and provide an assessment of outputs and outcomes from funded projects.

Our headline finding is that, based on analysis of the views of an independent panel of social scientists, **the Scheme successfully uses innovative assessment processes to identify and fund transformative research.** It achieves this not only through an innovative combination of assessment methods – elements of which will likely have wider utility in other funding schemes – but also through **a culture of learning, discussion, self-reflection and contestation, which has been allowed to develop around the Scheme.** Additionally, the outputs of the scheme are comparable in terms of both quality and quantity with similarly sized ESRC Standard grants, but more often appear to **present unusual inter-disciplinary perspectives, challenge established conventions of social scientific enquiry or have the potential to lead to a paradigm shift.** At the same time, projects from the Scheme tend to present **only the beginning of wider research agendas, rather than being closed and discreet projects with a fixed end-point.** We detail the various sub-components of this main finding in the pages below.

The Scheme in context

Several funding schemes exist both in the UK and abroad that seek to identify and fund transformative research. Most of these schemes either implicitly or explicitly result from a number of key considerations, including most notably:

- The need systemically to accept a certain degree of risk in terms of projects achieving successful outcomes, whilst also ensuring a means to control risk, especially in tough funding landscapes;
- The conservative bias of peer review processes, which tend to favour established methods, topics and individuals.

Responding to the challenges in identifying transformative research, assessment procedures of comparable schemes take many different forms, ranging from complete by-pass of peer review and de-risking through provision of very small grants, to large grants with flexibility for transformative endeavours, de-risked through extensive peer review and highly selective choice of candidates.

The ESRC Scheme offers a useful compromise in this respect. It uses a two-stage proposal assessment process with anonymous peer and panel review of applications to identify transformative scope followed by a 'Pitch-to-peers' further to assess quality and feasibility (de-risking).

Transformative research

Definition

In order to assess the overall efficacy of the Scheme in terms of its ability to identify and fund 'transformative research', it was first necessary to arrive at a suitable definition of the term. The ESRC aims to fund projects that present a methodological or theoretical innovation, application of existing

methods to new contexts, or unusual disciplinary or interdisciplinary perspectives. Moreover, the characterisation of high-risk, but with the possibility of high reward, for instance in the shape of a broad base of new knowledge, is a further supplementary definition in the documentation of the Scheme.

However, there are many further possible definitions and characteristics associated with transformative research: in the philosophy and sociology of science, the notion of paradigm shifts and displacement of old established methods or theories is noteworthy here. Other characteristics such as difficulty to find favour with peer reviewers, long lead times, or an association with early career researchers are noted in the wider literature on this concept. In the case of the social sciences, there is a further association between transformative research and scholars from various marginalised groups, e.g. along lines of gender or ethnicity.

To arrive at an expert definition of ‘transformative research’, we used a review exercise of 93 projects submitted to the Scheme (45 funded, 48 rejected at sift panel or Pitch-to-peers stage) and asked a pool of experts to rate each of them on a list of 17 criteria pertinent to transformative research, and to provide a verdict on whether or not they consider each proposal to be transformative. Analysis of results identified six characteristics that are strongly linked to reviewers’ overall judgements about whether projects are transformative, giving us a multi-faceted expert definition of the term:

- Pioneering theoretical or methodological innovation;
- Engagement of unusual disciplinary and interdisciplinary perspectives;
- High-risk research with the possibility of high reward;
- Research that could lead to a paradigm shift in its field;
- Research that is ethically sound;
- Research likely to require comparatively high amounts of funding to achieve successful results.

Success

Subsequently, a composite of our reviewers’ scores on these six criteria was found to be a good predictor of outcomes of the Scheme, defined in the simple binary of funded versus not funded. Taking the above six criteria and accepting that they apply to varying extent to individual proposals, **the Scheme successfully identifies and funds research with transformative characteristics**, as perceived by an independent panel of senior members of the community.

Profile and Attractiveness of the Scheme

Our review exercise showed that as a whole, the Scheme attracts submissions that present innovative theoretical or methodological innovations, whilst still having strong hallmarks of contemporary professionalised UK academia: suited to eventual publication in leading journals; ethically sound; and with scope for wider non-academic impact.

Individual applicants and institutions tend to view the Scheme favourably for a range of different reasons, notably in terms of presenting an avenue for creativity in what is perceived to be a conservative mainstream funding landscape, as well as having a relatively small grant Scheme to counter the growing emphasis on fewer, larger grants. However, there are currently two evident issues:

- From the institutional perspective, there is some difficulty in assessing which projects or candidates to put forward, as the aims of the Scheme are broadly defined and therefore unclear. Recruiting former applicants (successful or not) to the institutional internal selection processes has been a common way of tackling this issue, though this is an on-going learning curve;
- From the individual scholars’ perspective the Scheme has too much of a stand-alone character, with no obvious channels for follow-up funding or fast-tracking mechanisms towards larger grants in cases of successful results.

Function and operation of the assessment stages

Anonymous Review (Sift Panel)

The anonymous review of two-page outlines at the sift panel stage receives a high level of positive feedback from all stakeholders. Applicants in particular expressed support for the anonymity, whilst the two-page format also brings a range of advantages, including attractiveness to apply and developing abilities to succinctly distil and clarify research ideas. Besides lessening *a priori* advantages for established, senior scholars, the anonymity factor has also had some additional, unintended yet broadly positive effects: even senior researchers may have a lack of recent track record, for instance following a career break, or when wishing to change the focus of their research. The Scheme provided a possible 'way in' for such cases.

In terms of its function, **the sift panel stage is most clearly focused on identification of transformative research**: discussion at the panel meeting strongly focuses on this notion, and far less on other considerations (e.g. scholarly quality, risk, ethics, impact). The meaning of 'transformative', and whether or not a proposal can be judged as such, are key questions permeating these events. In particular, the involvement of individual panellists and panel chairs willing and able to encourage such debate are important success factors for this stage of the process.

Presentation Event (Pitch-to-peers)

The Pitch-to-peers presentation events have less of a clear focus on identification of transformative scope. Though this factor is still in play, **scholarly considerations, ability of the candidate and feasibility are more prominent areas of attention here**. At the same time, the interactive nature of these events gives candidates the opportunity to clarify issues that may have been left open by the short proposal reviews at the sift panel stage. The comparatively unusual nature of proposals submitted to this Scheme makes this space for exposition and questioning especially important.

Whilst the sift panel stage has broad appreciation from stakeholders, there are some concerns around the Pitch-to-peers events with regard to efficacy and equal opportunities. There are several acknowledgements that the format can encourage an excessively bullish, competitive and intimidating atmosphere, which suits some individuals better than others. Moreover, there is a perceived danger that presentation skills rather than substantive ideas are rewarded by this format. However, these criticisms are almost exclusively associated with the first call of the Scheme (2012-13). Subsequently, these dangers have been acknowledged and organisers and panellists have since managed to create an atmosphere and rules of engagement that appear to successfully address these issues.

Establishing a culture of transformative research

Throughout the evaluation, it was evident that the success of the Scheme has been dependent on a number of learning and feedback loops, and on allowing space and opportunity for discussion, debate and contestation. There has been a need for stakeholders to clarify, optimise and learn. This process has been fruitful: several **problems are evident when looking at the first call of the Scheme, but there is much evidence of improvement over time** on several factors, notably:

- Both the pool of applicants and eventual grant winners were dominated by male and senior scholars in the 2012-13 call; a significantly more balanced picture emerged in subsequent calls;
- Attendants' feedback on Pitch-to-peers events improved between the 2012-13 and 2013-14 events: overall organisation and enjoyment, as well as efficacy and quality of feedback were all judged higher for the second Pitch-to-peers;
- At the 2012-13 Pitch-to-peers fellow applicants judged presentations more harshly than assessment panellists did. In the subsequent two calls this harshness disappears and fellow applicants' scores are similar to or higher than those given by panellists;

- There was also significant disagreement in 2012-13 between panellists' and fellow applicants' scores at the level of individual applications, whilst in the most recent call, scores correlate strongly, indicating a greater common understanding of what the Scheme is intended to reward.

These improvements have not come about as a result of formal rule changes (there have been very few over the three calls). Instead, they are the result of learning and feedback, often implemented through the initiative of panellists and panel chairs. Many successful applicants have since also become panellists or advise their host institution on which proposals to put forward to the Scheme. The sum of these learning and feedback loops is critical in optimising the Scheme, and in developing an understanding and a culture of transformative research within the scheme and beyond, in eligible institutions and among the social science community more broadly.

Outputs, outcomes and future perspectives

In terms of productivity, the TR grants produce similar quantities of outputs as ESRC standard grants of comparable size and timeframe. Likewise in terms of quality, there is no apparent difference between Standard and TR grant outputs. In this sense, the vast majority of projects were successful at some level, albeit in many cases only as a proof-of-concept. There is little sense of undue risks or absent rewards.

On several criteria, the TR projects and individual outputs are judged more transformative than their counterparts in the Standard grants scheme, especially on three criteria:

- Engagement of unusual interdisciplinary perspectives
- Likelihood to challenge widespread assumptions of typical social scientific enquiry
- Potential to lead to a paradigm shift in its field

The TR grants also produce outputs that more often tend to suggest significant further work or follow-on funding may be necessary for results to come to full fruition.

However, this does not mean that TR grants categorically produce 'transformative' outputs while standard grants do not: many outputs in both categories were identified as transformative, though the TR grants considerably more so. On the criteria we measured, many outputs of Standard grants also have characteristics associated with being 'transformative'.

There are strong relationships between the transformative character and scope of the grant outputs, and their quality. Moreover, there is some evidence that outputs associated with early career researchers do less well in both these sets of dimensions. Such tendencies are to be expected, however, as early career stages might denote less experience in rapid production and publication of high-quality outputs.

Particularly for the more recent grants, results and critical outputs are still emerging. Even for the less recent grants, there are still 'publication tails' that are in progress. Many grant holders from both rounds analysed here have secured further funding from a range of different funders. Most grant holders are moreover either in the process of writing applications for further funding, have applications under review or are considering application in the near future. This mitigates to a large extent any concerns that might otherwise arise around the increased proportion of awards to ECRs in the later calls of the scheme: initial lack of experience to produce high quality outputs is likely to diminish as agendas progress. Relatedly, there is an overwhelming sense that funded projects are seen by grant holders as merely marking the beginning of a longer-term research agenda, rather than being discreet projects with fixed end-points. A long-term impact study in the future might therefore be an especially useful exercise.

There are very few projects that have fully or mostly failed to achieve successes. Several have so far seen only early, emerging findings, which is to be expected in what are often long-term endeavours. Those few that show no evidence of substantive outputs or successes appear to stem from over-ambitious research plans or elements thereof, which turned out not to be feasible. Such cases are extremely rare, compared with many ambitious plans that have in fact materialised.

Likewise, there are few projects where demonstrable and large-scale transformative impacts are already evident, due again to the proof-of-concept nature of many of the projects, and their consequent long-term perspectives. Where developments have occurred that may already be described as genuine substantive transformations, they are often around new types of collaborations. These might be interdisciplinary or between researchers and other groups (e.g. policymakers). Topical subjects combined with outreach activities in some cases also led to more rapid exposure than would otherwise be the case.

Summary of Recommendations

The ESRC has successfully created a space in which a growing culture and understanding of transformative research is taking hold. **The Scheme should continue operating in the future.** The wider effects of this emerging culture on the ESRC and social science community cannot be determined at this point, though an impact study in the future would likely also be a worthwhile undertaking in this regard.

Some operational aspects are worth modifying or clarifying, e.g. the rules of engagement at Pitch-to-peers events, how to manage conflicts of interest, how best to apply the ten-point scoring scale. However, most importantly **it is critical to maintain the open, collegial and self-reflexive nature of the Scheme.** Panellists need to be carefully selected, based at least to some extent on their willingness to engage in such self-reflection. Recruitment of former grant holders to the panel has provided especially helpful input, so this is worthwhile continuing and expanding.

It is worth considering introduction of clearer pathways to follow-up funding channels (or fast-tracking mechanisms to existing big grant schemes), to be clarified to applicants from the outset. Eligibility could be made dependent on a minimum profile of outputs and outcomes (e.g. 'proof of concept').

Whilst we see few concerns about the level of budget, increasing the length of allocated time of the grants (e.g. from 18 to 24 months) may be a useful step.

The anonymous review format has wide appeal. Extension to other schemes should be considered. However, a second step aimed at de-risking and checking applicants' scholarly ability is likely to be necessary. The two-page application format might likewise be useful elsewhere, especially in schemes that could benefit from the ability to mitigate some of the known deficiencies of peer review.

Introduction

This report presents an evaluation of the ESRC's Transformative Research Scheme (hereafter 'the Scheme') carried out by Technopolis between March 2015 and June 2016. The aims of this evaluation are as follows:

- To determine the extent to which the Scheme has funded research that is transformative;
- To analyse the operational and procedural elements of the Scheme;
- To inform future calls through the Scheme and other commissioning activity;
- To assess outcomes of the transformative grants.

The first three aims were covered in Phase 1 of the evaluation, which ran from March to September 2015, and the findings of Phase 1 are detailed in chapters 1 to 5 of this report. The last of the four evaluation aims was covered by Phase 2, which ran from January to June 2016; its findings are the subject matter of chapter 6.

Whilst this study is primarily a funding scheme evaluation concerned with operational and functional aspects, the notion of 'transformative research' itself is a complex one, with significance for broader questions about research funding and 'how science works'. This study has therefore necessitated engagement with these questions. Where relevant, we include findings about these issues. This report will therefore be of interest not only to stakeholders in the Scheme but also to the social scientific and research funding communities more broadly.

Our methodology for this evaluation is described in full in Appendix A. In brief, we draw our findings from:

- A desk research and literature review phase, which assessed the notion of transformative research itself with regard to its possible meanings and definitions, challenges of identifying and funding it, as well as its overall significance in the philosophy and sociology of science. This method component also contained a comparative dimension, establishing an overview of how other funding schemes have addressed the challenge of funding transformative research;
- A survey of successful and unsuccessful applicants to the Scheme (n=81);
- A total of 20 interviews (around 30 minutes in length each, conducted by telephone or Skype) with several different stakeholder groups:
 - Current grant holders (x6);
 - Assessment panellists and panel chairs (x5);
 - Research officers at grant holding institutions (x6);¹
 - ESRC officers charged with sponsoring or overseeing various parts of the scheme (x2);
 - Non-affiliated experts on research funding and transformative research (x3).
- Observation at the Sift Panel and Pitch-to-peers event of the 2014-15 call of the Scheme;
- A peer review exercise of 93 proposals submitted to the Scheme, where a pool of seven reviewers (all senior professors who have either acted as REF panellists or been in charge of their department's REF submission) scored each proposal on a range of criteria relating to the idea of 'transformative research'. Statistical analysis was subsequently carried out on the peer review data and its relationship to the outcomes of each proposal in the Scheme (of which reviewers had not been informed).

Our data collection and analysis for Phase 2 consisted of three main components:

¹ Some grant holders subsequently became panellists. We were especially interested in talking to these individuals as they have had multiple perspectives on the Scheme. Due to this overlap, the total number appears to exceed 20.

- Interviews with 20 grant holders (12 from the first call and eight from the second). Details of interviews are noted in Appendix I.
- Analysis of outputs and key findings data entered by grant holders into ResearchFish.
- A peer review exercise of 80 outputs (journal articles and working papers): 40 listed on ResearchFish as outputs from Transformative Research grants and 40 listed as outputs from ESRC standard response mode (hereafter: Standard) grants. Reviews were conducted by a panel of six peer reviewers (all of whom had already contributed to our Phase 1 review of applications), who, for each output, completed a template similar to the one used in our Phase 1 review: 15 characteristics pertaining to transformative scope, and an overall judgement on whether or not the output is transformative. Additionally, we added a REF-style quality-profiling matrix, and also asked reviewers to rate their own confidence in their judgement for each output. Reviewers were not told which outputs were from the TR scheme, or even that half of the outputs were not, though some disclosure was inevitable, as some outputs included an acknowledgement of the TR funding scheme. Details of the outputs review are noted in Appendix J.

Section 1 of this report introduces the Scheme itself and where it stands in relation to the key questions and funding challenges associated with transformative research, and also notes its similarities and differences to other comparator schemes. In section 2, we discuss our headline conclusion on the overall efficacy of the Scheme and the meaning of transformative research within it. Section 3 will look at the various sub-components of these findings, by presenting conclusions on the various assessment stages of the Scheme, their respective roles, importance and operational aspects. Section 4 will reflect in more depth on the wider learning curve that this scheme has constituted for its various stakeholders. In section 5 we note further conclusions about the perception, operation and attractiveness of the Scheme. Finally, section 6 presents our findings on the outputs, outcomes and longer-term future trajectories of funded projects, before making a set of recommendations in the concluding section.

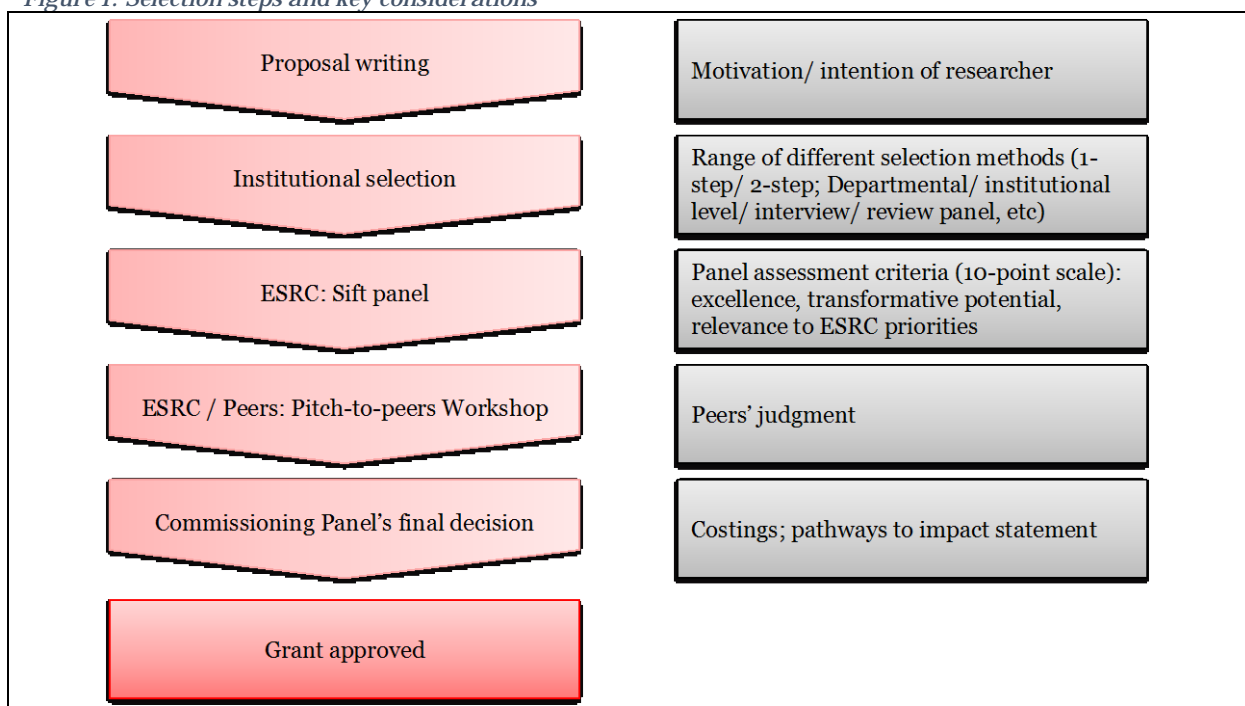
1 The Scheme: description, context and comparison

1.1 The ESRC scheme

ESRC launched the Transformative research Scheme in 2012, with three calls for proposals (2012/13, 2013/14, 2014/15), responding to widely held concerns that conventional research funding tools do not reliably identify and fund research with transformative scope. Each successful project received around £200,000. In the 2012/13 call there was an additional payment to the principal investigators' (PI's) institution of £50,000, to be used for promoting and encouraging further transformative research within the social sciences.

There are several steps involved up to the decision of whether a project is funded, summarised in Figure 1.

Figure 1: Selection steps and key considerations



- The institutional level:** the prospective PI submits their proposal to their institution, which then submits to ESRC. The limits on the number of applications and the eligibility of institutions varied across the three calls:

 - 2012/13: only UK research organisations that received more than £100,000 of ESRC research funding in 2011/12 could submit to this scheme. Only **one** application was accepted from each eligible research organisation, except those that had received in excess of £3 million of research funding, as detailed in the 2011/12 ESRC Annual Report, who were allowed to submit **two** proposals;
 - 2013/14: any UK research organisation eligible for Research Council funding could submit **one** proposal, except those in the 'Top 11 Institutions by Research Expenditure' in 2012/13, as detailed in the 2012/13 ESRC Annual Report, which were allowed to submit **two** proposals;
 - 2014/15: any research organisation eligible for Research Council funding was allowed to submit **two** proposals.
- Sift panel:** once proposals have been submitted to ESRC, the next stage consists of a panel review of submitted applications. Applications are submitted to the panel **anonymously** and consist of a 2-page outline of the proposed research. The panel judges applications on excellence,

transformative quality as well as impact and relevance to ESRC priorities. Agreement is reached after reviewers grade applications on a 10-point scale, which is also used for the subsequent stage of the assessment process:

Table 1: Sift panel – assessment criteria

Grade	Comments
10	The proposal is exceptional in terms of its potential transformative nature and scientific excellence with impact and fit to the aims of the call. Should definitely be shortlisted on the basis of the proposed work.
9	The proposal is outstanding in terms of its potential transformative nature scientific excellence with impact and fit to the aims of the call. Should definitely be shortlisted on the basis of the proposed work.
8	The proposal is excellent in terms of its potential transformative nature and scientific excellence with impact and fit to the aims of the call. Should be shortlisted as a priority but would benefit from addressing the feedback from assessors in the full proposal
7*	The proposal is very good in terms of its potential transformative nature and scientific excellence with impact and fit to the aims of the call. Should be shortlisted as a priority but will need additional questioning at the Pitch to Peers workshop.
6	The proposal is good in terms of its potential transformative nature and scientific excellence with impact and fit to the aims of the call. Could be shortlisted but is not of a consistently high quality to be a priority for shortlisting. Will require additional questioning at the Pitch to Peers workshop.
5	The proposal is good in terms of its potential transformative nature and scientific excellence with impact, and its fit to the aims of the call. Could be shortlisted but the proposal would require some substantial questioning at the Pitch to Peers workshop and potential modification when the full proposal is submitted.
4	The proposal is good in terms of its potential scientific excellence with impact but does not entirely meet the aims of the call in terms of being genuinely transformative in nature. Could be shortlisted but is not of a consistently high quality to be a priority. Will require some substantial questioning at the Pitch to Peers workshop and potential modification when the full proposal is submitted.
3	The proposal is satisfactory in terms of its potential scientific excellence with impact but where, if shortlisted, the proposal would need some major amendments to ensure it was transformative in nature.
2	The proposal is satisfactory in terms of its potential scientific excellence with impact but does not fit with the aims of the call in terms of being transformative in nature.
1	This is a reject grade. A reject grade should be awarded to a proposal which is flawed in its scientific approach, or is repetitious of other work, or otherwise judged not worth pursuing; or which, though possibly having sound objectives, appears seriously defective in its methodology.

*Minimum pass mark

- **Pitch-to-peers workshop:** Shortlisted applicants are invited to attend a ‘Pitch to Peers’ workshop. At this event each shortlisted applicant presents their proposal, which is assessed by other shortlisted applicants and the assessment panel. This event includes evening activities and overnight accommodation. At the workshop, applicants are divided at random into groups of around 10-12 people, including some Commissioning Panel members. Applicants present their proposal to the group, and the group then questions and grades the proposal;
- **Final Decision:** scores from the workshop are then analysed by the assessment panel, which makes a final recommendation on the grants to be funded. Funding decisions are made immediately following the workshop and communicated to applicants shortly thereafter. The final costings and the ‘Pathway to Impact’ document are particular additional foci for the Commissioning Panel prior to the grant being awarded.

Table 2: Details of applications

	2012/13 call	2013/14 call	2014/15 call	Scheme totals
Total applications	67	69	106	242
Shortlisted for Pitch-to-peer workshop	32	26	25	83
Funded	20	13	12	45

1.2 What is ‘transformative research’?

The idea of transformative research originates in the philosophy of science, specifically in Kuhn’s book *The Structure of Scientific Revolutions* (1970). This proposed a definition of science not based on method but on the idea that science is what scientists do. It described two kinds of science: ‘normal’ science, which is the everyday business of incrementally improving existing theories; and ‘revolutionary’ science, which involves the overthrow of existing theories or ‘paradigms’ and their replacement by new ones. Scientific revolutions tend, according to Kuhn, to be opposed by the scientific establishment and revolutions often only take hold when there is a generational shift within the scientific community. This meant it could be hard for revolutionary science to be published or funded since the establishment also provides scientific gatekeepers via peer review.

These ideas have since been developed further, making explicit connections to funding mechanisms and underpinning the conventional wisdom among research funders that peer review is inherently conservative (Wessely 1998; Horrobin 1996; Roy 1985; Lakatos & Musgrave 1970).

Luukkono et al (2015) note that this interplay between defence of existing paradigms and new transformative research is still apt today, but that there are instances of transformative research that are not fully symptomatic of this tension: unexpected discovery of new methods or fields do not necessarily present direct challenges to existing paradigms, and thus weaken the notion of conservatism as a barrier to transformative research. Especially where such new discoveries have a high public profile, the need for engagement and a certain degree of requirement for new approaches is more likely to be accepted. Nevertheless, such instances can likewise be disruptive to established paradigms, theories or practices.

The ESRC defines transformative research as “involving pioneering theoretical and methodological innovation, the novel application of theory and methods in new contexts, and/or research based on the engagement of unusual disciplinary and interdisciplinary perspectives.”² This is one of many possible definitions, though many comparator schemes (see Section 1.3) Use similar definitions. Arguably, this is broader than Kuhn’s ‘revolutionary science’ but it still maintains the idea of working outside the established conventions and having potential to be disruptive within the world of science and carries with it the idea that it challenges current thinking.

Much transformative research funding is in the hard sciences, especially in the life sciences and medicine. Transferring the idea of transformative science from the ‘hard’ to the social sciences is a bold experiment on the part of the ESRC. But whatever the context, transformative research is hard to identify, especially ex ante. Inherently, it is not only outside the mainstream but has the potential to cause change – and while the former is relatively easy to identify, the change potential is more difficult.

These ideas started to have significant influence on funding, especially after from the 2000s onwards. Particularly influential was debate in the US scientific community about how to fund the high-risk, often interdisciplinary work that would lead to scientific revolutions – or, in the terminology then adopted, ‘transformative’ research (National Science Board 2007). In practice NSF had been trying to

² from the Terms of Reference

fund transformative research – “high-risk, high-reward research that might not pass the traditional peer review process” – at least since 1990. Wagner and Alexander note that this funding tool was under-utilised, indicating internal unwillingness or inability in NSF to tackle risk (Wagner & Alexander 2013).

1.2.1 Two notions of transformative research: transforming science or transforming society

It is worth reflecting briefly on two fundamentally distinct notions of ‘transformative research’: the first is, as outlined above, research that seeks to pioneer new approaches and thus transform the way science operates. The second is specific to the social sciences and is research that aims to trigger societal changes, either within a research project or as a result of it. This kind of approach and its evaluation is discussed by Mertens (2008) but it is not the central subject of this study: in the ESRC’s Scheme, and therefore in this evaluation, ‘transformative’ refers to the nature of the research or science itself. Nevertheless, it is important briefly to consider the connection between the two.

There is no evidence of a direct link between research being transformative of society and being transformative of scientific practice. In fact, many examples easily come to mind of ‘normal science’ having profound effects on world events (e.g. vaccine development in the recent Ebola outbreak, which used old-fashioned isolation techniques). Nevertheless, many transformative research funding schemes aim at some level not only to fund research that is transformative in terms of its scientific implications, but also in terms of societal significance (most of our comparator schemes in Section 1.3 contain an ‘impact’ dimension at some level in their criteria).

As there is no direct link between the two, efforts in the Scheme to identify transformative scope in scientific terms does not readily imply that merits in that domain will bring about societal transformation. Instead, there is besides the criteria discussed here also a requirement for proposals submitted to the scheme to add a statement of impact as part of the two-page outline proposal. These impact dimensions are discussed at the sift panel meeting, and are also to be considered in the assessors’ final judgement following the Pitch-to-peers workshops. Additionally, the subject matter of proposals must be relevant to one or more of the ESRC’s strategic priorities:

- Economic Performance and Sustainable Growth;
- A Fair and Vibrant Society;
- Influencing Behaviour and Informing Interventions.

Table 3: 2012/13 call submissions – fit with ESRC strategic priorities

Priority	Number	%
Economic Performance and Sustainable Growth only AND with one or more other priorities	25	38
A Fair and Vibrant Society only AND with one or more other priorities	34	52
Influencing Behaviour and Informing Interventions only AND with one or more other priorities	60	91
Economic Performance and Sustainable Growth ONLY	2	3
A Fair and Vibrant Society ONLY	2	3
Influencing Behaviour and Informing Interventions ONLY	20	30

Source: ESRC

Whilst most of the details of the assessment system for the Scheme are related to transformative elements in purely scientific terms, this additional dimension provides a way of ensuring projects have some degree of focus on key societal issues and challenges. Given the lack of linkage between these two elements, the Scheme implicitly involves a balancing act between these two notions of transformation. This ‘impact’ dimension features alongside a range of other considerations, which need to be addressed alongside transformative scope in the assessment process (notably scholarly quality,

feasibility and ethics). As we show in Sections 2 and 0, the question is not so much whether the Scheme considers *only* transformative scope, but whether transformative scope is visibly rewarded by the Scheme despite the need in contemporary professionalised academia to also assess for quality, feasibility, ethics and – not least – wider impact.

1.3 Funding transformative research: problems and solutions

Research funders generally experience that it is easy to **aim** to fund high-risk/high reward or transformative research but that it is in practice difficult to devise mechanisms for doing so. Some decide to ‘mainstream’ risky research but this may not be very successful, because the main review criteria tend to squeeze out transformative, interdisciplinary or risky work (Arnold et al. 2013; Häyrynen 2007). Increasingly funders provide ring-fenced funding for transformative research, in recognition of the difficulties it may have in getting through conventional peer review. The same logic is often used for young researchers, taking them out of a competition for resources with senior members of the scientific community that they would be unlikely to win. Thus funding schemes targeting transformative research exist across a range of different countries and disciplines. Table 4 summarises a selection of transformative research funding schemes considered here in order to establish common practice as well as different approaches that exist in this area (details of these schemes are listed in Appendix D).

Table 4: Overview of other transformative research funding schemes

Organisation (Country)	Name of scheme
EPSRC (UK)	Bright IDEAS Award
ERC (EU)	Starting grants and Advanced grants
NIH (USA)	Director’s Pioneer Award Programme
NSF (USA)	NSF Small Grants for Exploratory Research (SGER)
RCN (Norway)	Young Research Talents
Wellcome Trust (UK)	Sir Henry Wellcome Commemorative Awards for Innovative Research
NSERC (Canada)	Discovery Frontiers

In general terms, there are several acknowledged difficulties in the task of funding transformative research, with multiple different ways of responding to them. Two related challenges are especially worth considering: establishing a risk-taking culture in the funding landscape, and tackling the potential for conservatism among peer reviewers. Underlying these challenges is a further question that we turn to subsequently: in order to encourage transformative research and manage risk, should a funding organisation emphasise scrutiny of proposed projects, or of the individuals proposing them?

1.3.1 Risk and conservatism

In the task of ensuring the efficiency and accountability of funding bodies, peer reviewers may be reluctant to fund research entailing a higher-than-usual degree of risk, even at the expense of potential high rewards (Chubin & Hackett 1990).

Across transformative research funding schemes there is the acknowledgement that transformative research carries a higher than usual degree of risk. It is important to note however that ‘risk’ has a number of different dimensions in research, with a key distinction worth noting here:

- In these schemes, ‘risk’ refers to the heightened uncertainty of outcome: where the scope of research is especially ambitious and involved some level of uncharted territory, the likelihood of successful results diminishes;
- Operational risks relate the safety of the researcher or the research participants. These are generally not addressed differently in transformative research schemes, as they relate to fundamental guidelines of research ethics.³

This higher level of risk associated with transformative research is widely acknowledged, and most funding schemes for transformative research state openly that a higher degree of risk in terms of possibility of failure is expected.

The possible presence of conservatism in the peer review process may also systematically disadvantage research challenging established paradigms. Häyrynen (2007) notes that much of the capacity to enable transformative research stems from fundamental attitudes across the community of researchers, reviewers and funders. Table 5 contrasts the two possible extremes in this respect.

Table 5: Extreme opposite mind-sets to risk-taking

	Researcher	Reviewer	Funder
High ambitions	Applies unconventional ideas and aims for significant results	Rewards innovative research plans and accepts shortcomings in proposals	Looks to achieve scientific breakthroughs with funding and tolerates failures, aims for dynamic science and research
Fear of risks	Conservative choice of subject designed to secure funding	Emphasises scientific quality, merit and viability	Aims for certain, measurable results, strategic goals and predictable changes

Source: Häyrynen (2007)

At the level of decision-makers, there is broad consensus that a higher level of risk than normal has to be accepted if transformative research is to be funded. At the level of the researcher, there are some considerations that may affect their readiness to take risks. It has been noted that the interdisciplinary and heterodox research are discouraged by the UK’s research assessment system – the REF – and that more generally, researchers tend to favour ‘safe’ options more likely to produce excellent research conducive to career advancement (Häyrynen 2007: 21-22). At the level of reviewers, this framework suggests that it may be necessary to steer towards risk-taking through explicit inclusion of transformative scope and character, or high-risk/ high reward content in the assessment criteria.

The selection of schemes noted here highlights a range of different approaches to successfully funding transformative research.

- The EPSRC’s IDEAS scheme – which in part inspired the ESRC’s Scheme – is similar in terms of reducing the role of peer review and instead involving a presentation event;
- Many other schemes have opted for a reduced role for peer review, though this is operationalized differently. For instance, NSF’s SGER scheme bypasses the notion of peer review altogether, leaving decisions on grants to the discretion of the programme officers. This approach self-evidently involves significant trust in the officer’s judgement;
- The Wellcome Trust’s Commemorative Awards Scheme takes a different approach: like the ESRC’s Scheme it has a two-step process, where the first round likewise rests on a two-page outline proposal, whilst the latter is a full peer review exercise. As such, the role of peer review is not so much diluted, but the first stage – though also peer reviewed – has as chief criteria to judge the innovative character of proposals. This ensures that all full applications moving to the peer review

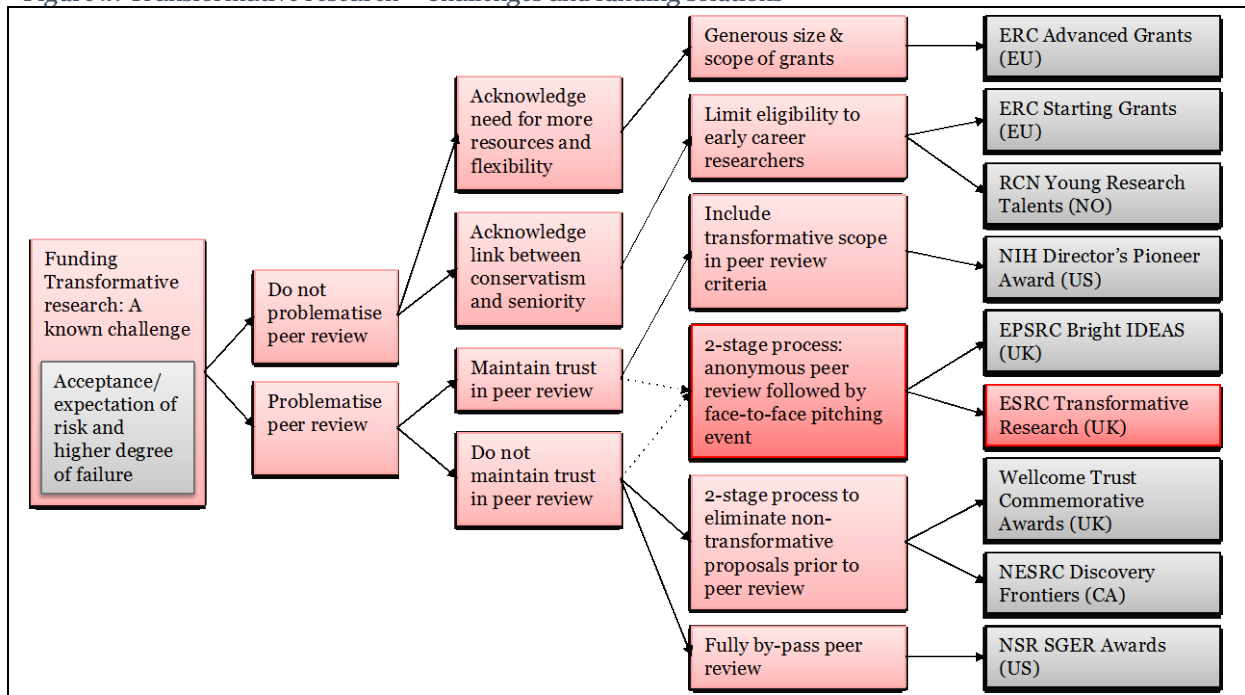
³ The NSF’s SGER scheme includes potential risk to the PI in its criteria, though it is unclear to what extent such risk is accepted or viewed as part of transformative research projects (Häyrynen 2007).

round will have some degree of transformative character, mitigating the conservative effects of peer review by ensuring a pool of exclusively transformative proposals from the start;

- The NIH’s Director’s Pioneer Award programme takes perhaps the most straightforward approach, by opting for a typical peer review exercise, but including the need for innovative scope in the assessment criteria. The underlying assumption is that peer review is not necessarily and inherently conservative, if peers are under instruction not to be, and have explicit criteria related to transformative scope.

Figure 2 schematises the different approaches taken by comparator schemes in relation to peer review, conservatism and the need for the right level of risk avoidance, ranging from large grants underpinned by scrutiny over the individuals applying for them (ERC Advanced grants), to very small grants where peer review is fully by-passed (NSR SGER awards). The ESRC’s Scheme offers a compromise solution, not veering too far towards either extreme. Fundamentally, peer reviewers lead the selection process, and there is implicit trust that they will be able to identify and fund transformative research proposals, through the existence of assessment criteria emphasising these qualities. However, by adding the anonymity at the first assessment stage, and by including fellow applicants in the scoring at the Pitch-to-peers events (see Sections 3.2 and 3.3) some devices are put in place, which are designed to potentially address some of the known hazards around peer review.

Figure 2: Transformative research – Challenges and funding solutions



Source: Technopolis

1.3.2 Review panels and remote peer review: expertise and strategy

Luukkonen et al (2015) furthermore point out that there is a key distinction in peer review, namely between use of remote peer reviewers and review panels. The former are generally selected based on specialist expertise, i.e. they are the best-suited available individuals capable of fully understanding and judging the merit of particular proposals. Review panels by contrast tend to be composed of generalists: a broad spectrum of expertise allows panels to make a reasonably well-informed judgement on a larger number of proposals. Furthermore, review panels tend to have a clearer understanding of the organisational context in which their reviewing activities are taking place. This means that whilst remote reviewers are relied on to provide the best possible academic judgement, it is accepted that they can conduct their review without knowledge of the organisational or policy context

in which the review exercise is taking place. For panel reviews, criteria such as a proposal's 'fit' to the funding scheme – which inevitably involves a degree of contextual understanding – tend to play a role. This suggests that panel review is better suited to transformative research funding schemes, and that indeed, comparator schemes in their research tend to opt for panels rather than remote reviewers.

1.3.3 *Funding people or funding projects*

In recent years, there has been a rise in project-based funding (as opposed to e.g. institutional funding). Though this generally ensures that research budgets are spent on carefully planned endeavours, Luukkonen et al (2015) argue that this very notion of careful planning limits the flexibility of researchers and reduces the capacity for serendipitous new ideas, or to follow through on these.

Laudel and Gläser (2014) note that transformative research is significantly harder to plan in advance than research projects conducted within the confines of existing topics and methodologies. The inevitable uncertainty attached to transformative research means that parameters such as time, financial resources or precise details of projected outcomes are harder to set. Due to its novelty and potentially multiple un-tested dimensions, transformative research may require more flexible timeframes and resources. Budgeting and timetabling may be done robustly when there is substantial existing knowledge about the chosen methods and fields of application, but in the absence of such knowledge, parameters are harder to set (Laudel & Gläser 2014).

The difficulty of planning transformative research as pre-scripted projects has led some targeted schemes to fund people rather than projects, most notably in ERC grants (Luukkonen et al 2015), which aim to provide large grants to individuals or teams with a strong track record, allowing for flexibility that can accommodate for activities not readily expressible as clearly defined and costed projects. Such approaches naturally place heavy emphasis on the qualifications, track record and experience of the PI or the research team, aiming to provide funding to those most likely to produce good results.

However, several authors have highlighted that peer review tends to favour researchers from dominant groups. This is the case for gender and ethnicity, and especially in relation to seniority (see Roy 1985; Travis & Collins 1991). A problem directly opposed to that of long lead times and uncertain outcomes of transformative research is that transformative research is often associated with marginalised groups who are typically disadvantaged by academic selection processes, e.g. early career researchers and female researchers. In line with early contributions by Kuhn (1970), those likely to be situated in a current scientific paradigm and most prone to defend it are researchers who themselves established said paradigm and became senior authorities in their field because of it.

In the case of the social sciences, this issue becomes especially amplified, as many of the key transformative, paradigm-shifting developments of the last several decades have come from groups other than senior, established, white, male researchers (e.g. feminist methodologies, post-colonialism, queer theory, intersectionality). Feminist research methods are one of the clearest examples of paradigm shifting research in the social sciences (see e.g. Oakley 1998), whilst the disadvantage experienced by female researchers in fields of academic life ranging from career progression to citation counts is among the most clearly documented (Lariviere et al 2013).

Whilst especially in the social sciences there is a degree of association between transformative research and less dominant demographic groups in the scientific community, the connection is not absolute: senior (as well as white, male, middle-class, able bodied) researchers are also principally capable of transformative endeavours. As such, some transformative research funding schemes have elements of preference for non-dominant groups, but these tend to be notional. The NIH's scheme for instance has a brief note specifically encouraging individuals from underrepresented groups to apply. However, this is a standard element in many funding schemes (including non-transformative), and there is no evidence of a heightened effort or more substantive measures to facilitate equal opportunities in transformative research funding schemes.

In terms of seniority, the ERC's starting grants highlight an exception to this, as these are specifically reserved for researchers in the early stages of their career. As these grants are large and cover a long period of time in contrast to other comparator schemes listed above, they effectively give early career researchers the resources to establish themselves in their field. Whilst it is therefore important that funding schemes for transformative research are not prohibitive to non-dominant groups, few schemes opt for a systematised focus on them. Broader mainstreaming of these groups (e.g. representation among peer reviewers) appears to be an appropriate course of action.

Whilst on the issue of peer review and risk management the ESRC's Scheme offers a compromise solution in the context of comparator schemes, on the question of funding people or funding projects, it falls squarely into the latter category. Although efforts have also been made to widen representativity of panellists (highlighted by our interviews), the focus on projects is made especially clear in the first assessment round, where details of applicants are kept fully secret from the assessors.

Neither approach – funding people or funding projects – is necessarily superior. In a tough wider research funding landscape, methods of funding transformative research through comparatively small grants – i.e. funding project – is an obvious approach to take. Moreover, given the particularly evident link between transformative research and non-dominant groups in the social sciences (feminism, postcolonialism, etc.), opting for anonymity of applicants becomes an especially suitable approach.

The analysis here has situated the Scheme in the wider context of both the concept of transformative research, as well as the wider landscape of schemes attempting to identify and fund it. The following chapters will consider whether and how it is actually a successful tool to identify and fund transformative research.

2.1.1 Towards an expert definition

Beyond the input from applicants, we sought to establish a definition of ‘transformative research’ in a more systematic way, using our project review exercise. Drawing on the ESRC’s own definition, as well as our literature review, we note a list of 17 characteristics and definitions of transformative research:

1. Research that presents a pioneering theoretical or methodological innovation (ESRC headline definition)
2. Research that presents a novel application of theory or methods to a new context (ESRC headline definition)
3. Research that is based on the engagement of unusual disciplinary and interdisciplinary perspectives (ESRC headline definition)
4. Research likely to involve an unusually high degree of risk in terms of likelihood to achieve successful research results (Acknowledged by ESRC and across comparator schemes)
5. Research likely to involve an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants (Häyrynen 2007)
6. Research that can be described as high-risk, but with the possibility of high reward (ESRC supplementary definition)
7. Research that challenges widespread assumptions of typical social scientific enquiry (related to Kuhn 1970 & noted in our stakeholder consultations)
8. Research that could lead to a paradigm shift in its field (Kuhn 1970)
9. Research likely to raise considerable interest in the social scientific community (ESRC supplementary definitions & stakeholder consultations)
10. Research that might struggle to suit the scope of the leading journals in the social sciences (based on Kuhn & literature on the conservatism of peer review)
11. Research that may struggle to find favour with peer reviewers in its aligned disciplines (Horrobin 1996; Roy 1985; Wessely 1998; Luukkonen 2012)
12. Research that has particular capacity for eventual application of results (if successful) and societal transformations outside of academia (Mertens 2008)
13. Research likely to require comparatively long lead times to achieve successful results (Laudel & Gläser 2014)
14. Research more typical of early career researchers than of senior, established scholars (often included in rationale for early career researcher grants, e.g. ERC starter grants)
15. Research that, even if successful, may struggle to gain widespread acceptance in the academic community (Kuhn 1970; Horrobin 1996; Roy 1985; Wessely 1998; Luukkonen 2012)
16. Research likely to create a broad base of new knowledge and insight (ESRC supplementary definition)
17. Research likely to require comparatively high amounts of funding to achieve successful results (Laudel & Gläser 2014)

The main aim of our project review exercise was to understand how research proposals funded under this Scheme are transformative. The above list of potential criteria was the basis for the review template (See Appendix A.5.2) and reviewers were asked to score each of 93 proposals submitted to the Scheme on a five-point likert scale for each of the 17 possible characteristics. In addition, our reviewers were also asked for a final verdict, i.e. ‘would you describe this proposal as ‘transformative’ (yes/ No/ Cannot say)? Full description of our statistical analysis can be found in Appendix C. Our main aims in this exercise were:

- To assess, which of the 17 possible characteristics had the greatest influence on reviewers’ final verdicts, i.e. which particular high or low scores acted as predictors of whether overall a proposal

was deemed transformative or not. Effectively, this would provide a systematised external expert definition of transformative research;

- To cross-examine reviewers' scores and actual outcomes in the Scheme to identify the characteristics of proposals likely to lead to success. Specific outcome categories we considered were:
 - Rejected at sift panel stage;
 - Rejected at Pitch-to-peers stage;
 - Funded;
 - Scores out of ten given at the Sift Panel;
 - Scores out of ten given at the Pitch-to-peers event;
- To assess the extent to which proposals that were ultimately successful in the Scheme reflected particular strengths in terms of those features highlighted by reviewers as being notably markers of transformative research.

2.2 Main findings from the project review exercise

2.2.1 Defining 'transformative research'

Addressing our first aim, i.e. to understand which factors most influenced our reviewers in their final verdict, analysis of the review data found that verdicts were predicted by the scores they gave on the following criteria,⁴ providing us with an expert-led definition of transformative research in the social sciences:

- Research that presents a pioneering theoretical or methodological innovation
- Research that is based on the engagement of unusual disciplinary and interdisciplinary perspectives
- Research that can be described as high-risk, but with the possibility of high reward
- Research that could lead to a paradigm shift in its field
- Research likely to involve an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants (Negative)⁵
- Research likely to require comparatively high amounts of funding to achieve successful results

2.2.2 The effectiveness of the Scheme

Using a composite indicator deriving from the six characteristics noted above, we find that our reviewers' scores are a strong predictor of outcomes of the Scheme (Table 6).

In other words: whether or not projects are funded by the Scheme is strongly influenced by the above characteristics, though no single one of them is an adequate predictor on its own. The Scheme is therefore successful at identifying and funding transformative research (as understood by our experts), as long as a pluralistic definition of the term is accepted, constituting variable combinations of the six features set out above.

Based on a simple probabilistic model, we find our composite score predicts respective outcomes of the sift panel and Pitch-to-peers events (though not quite as clearly the overall Scheme outcomes, see Appendix C). In this sense, both assessment steps add value to the overall Scheme outcomes, although a more nuanced picture is revealed by the detailed analysis of the various assessment steps in Section 3.

⁴ This means that each of these scores are significant predictors of our reviewers' the final verdict. This also implies that we did not find a (statistically) significant association between the verdict and the other 11 criteria.

⁵ i.e. a low score on this criterion was associated with a positive final verdict

Table 6: Reviewers’ score as predictor of outcomes of the Scheme

	Overall scheme outcomes Indicator 1: The project has been accepted, and zero otherwise	Sift panel outcomes Indicator 2: The project has been accepted or shortlisted, and zero otherwise	Pitch-to-peers outcomes Indicator 3: The project has been accepted, zero if it was shortlisted
Reviewers’ score (Composite indicator)	0.381*	0.260*	0.344*
	(0.114)	(0.117)	(0.135)
Number of observations	173	173	127 (excludes <u>rejected</u> projects)

Standard errors in parentheses. The (*) in each cell means that the composite indicator is statistically significant at 95% confidence level.

2.2.3 A pluralistic process for a pluralistic concept

Whilst these findings suggest that the Scheme as a whole as well as each of the main assessment stages successfully identifies transformative research as defined above, there is considerable variation at a more granular level, notably between the two assessment stages. Though the overall efficacy of the assessment process might initially suggest a multi-step, yet essentially linear path towards positive outcomes, this is in fact not the case. A critical point to note is the weak association between projects’ sift panel scores and their subsequent scores at Pitch-to-peers events (Table 7). This weak relationship suggests a high level of competing perspectives, understandings and disagreements within the assessment process.

Table 7: Association between sift panel and Pitch-to-peers scores

	Strength of the relationship	R-squared (% of the variance explained by the model)	Number of observations
All proposals	Significant ⁶	6%	69
2012-13 call	No-significant	0%	28
2013-14 call	Significant	21%	21
2014-15 call	No-significant	0%	20

Data gathered as part of the statistical analysis of the project review exercise. Analysis based on univariate regression analysis.

This indicates that the very mixture and plurality of the assessment process has facilitated this overall impressive result. As the following sections will demonstrate, the Scheme offers space to debate the meaning and identification of transformative research, with assessment panellists, panel chairs and applicants able to provide feedback and engage in discussions. The meaning and identification of transformative research is problematised, discussed and disagreed upon regularly. Our observation data presented in Section 3.2.1 demonstrate this especially clearly.

Briefly looking ahead, in subsequent sections we present additional data supporting the idea that a pluralistic system including discussion at multiple points is an important hallmark of the Scheme that decisively impacts its success in identifying transformative research. The inclusion of past applicants in the assessment panels, and the evidence of continuous learning in the Scheme suggest that in time this capacity to identify and fund transformative research might well become even stronger and perhaps more clearly defined. More important, it underscores the importance of selecting the right individuals to contribute to the assessment procedure: interest in contributing to such a learning exercise, and willingness to debate the complex notion of ‘transformative research’ are critical characteristics.

⁶ The significant association found across all proposals seems to be driven by the results in the 2013-14 call.

3 The assessment stages: individual significance

In this section we present findings relating to the function and efficacy of each of the assessment stages of the Scheme. This includes foremost the sift panel stage and the subsequent Pitch-to-peers stage. However, prior to submission to the Scheme, there is the stage of institutional selection: at this point, we can learn what kind of proposals are submitted to the Scheme in the first place, i.e. whether it attracts proposals with the characteristics hoped for by the ESRC, and what kind of mechanisms are in place to facilitate identification of such proposals at the institutional level.

3.1 Stage 1: The institutional level

ESRC data suggest that institutions have to select applications to put forward to the Scheme out of a total of between 2 and 12 internal proposals per institution, per call. Institutions submitting proposals to ESRC used a variety of different selection processes, including panel reviews, interviews, 1-step and 2-step processes, as well as various combinations of department-wide and institution-wide approaches, and also anonymised and non-anonymised applications. The following is known about these different approaches:

- Institutions adopting a 2-stage process were more successful;
- Institutions adopting an anonymised internal selection process were more successful;
- Panel meetings as opposed to interview panels tend to generate the most successful proposals;
- One institution chose to use a Pitch-to-peers internal selection process. This resulted in a successful application;
- Those institutions that opened up the opportunity to apply to all across the institution generated successful proposals. None of those that chose to approach selected departments or researchers were successful;
- Institutions adopting a department-led approach to the *first stage* of their institution wide selection process had a higher success rate.⁷

Table 8: Outcomes of institutional selection processes

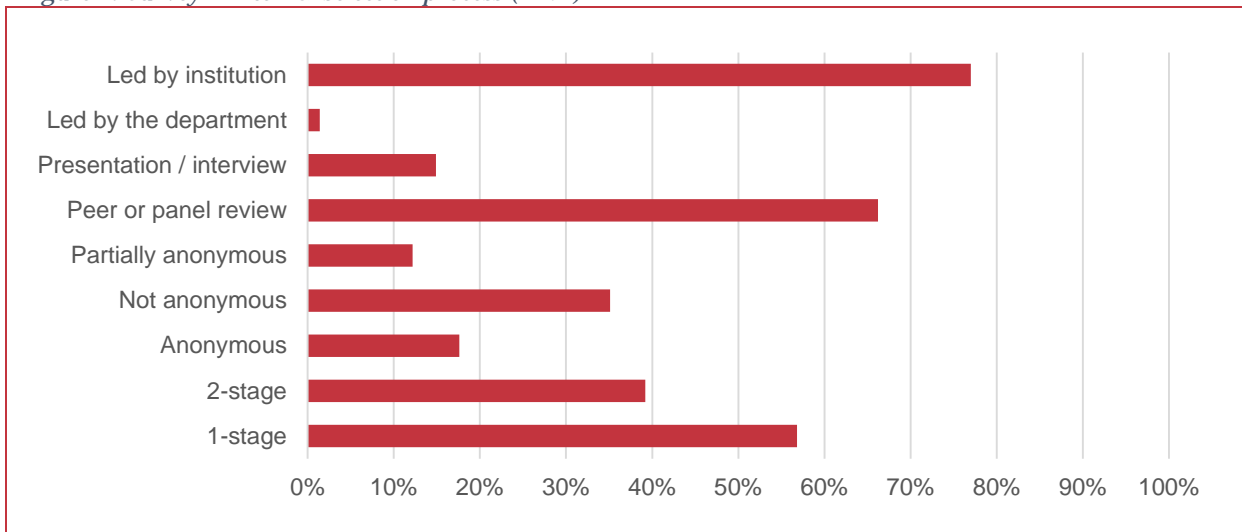
Selection methods	Total		Successful		Shortlisted but not successful		Unsuccessful		Office reject	
1 stage	27	46%	4	15%	9	33%	17	63%	1	4%
2 stage	32	54%	7	22%	13	41%	19	59%	0	0%
panel meeting used	35	59%	7	20%	7	20%	20	57%	1	3%
interview used	7	12%	1	14%	0	0%	6	86%	0	0%
university wide	38	64%	7	18%	7	18%	24	63%	0	0%
department led	13	22%	4	31%	2	15%	7	54%	0	0%

Source: ESRC internal analysis (N.B.: of those who submitted data, 2013/14 call only)

Our survey data reflect the above numbers: institutions employ a range of different techniques for internal selection, though from the applicants' point of view it is overwhelmingly the institutions rather than the departments leading the process.

⁷ ESRC internal analysis – does not cover the most recent call.

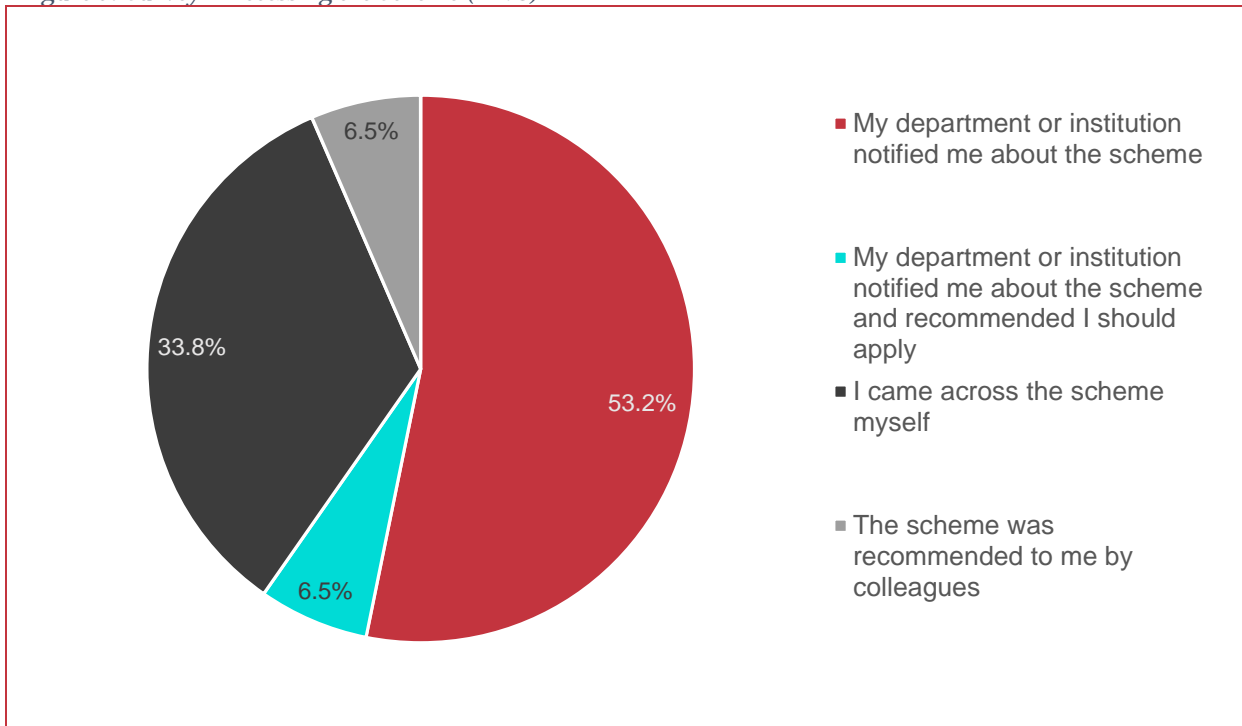
Figure 4: Survey – Internal selection process (n=74)



N.B.: respondents were invited to select as many or as few options as they wished, hence intuitively related options do not always add up to 100%

Our interviews furthermore highlighted that there was a somewhat experimental approach taken by most institutions, especially in the earliest call. They were unsure what the Scheme was looking for (i.e. what might constitute a promising proposal), approaches ranged from copying the Scheme’s assessment process (i.e. two-stage, with anonymity ensured for the first of these), to much simpler selection comparable to other funding application processes. We do find evidence in our survey that whilst institutions notified applicants about the Scheme, applicants generally decided to apply on their own initiative, rather than being explicitly asked by their institution.

Figure 5: Survey – Accessing the Scheme (n=78)



Institutions often draw on individuals associated with the Scheme in order to understand better which proposals to submit. Several interviewees noted that past applicants who were shortlisted to a Pitch-

to-peers workshop (whether successful or not) were asked by their institution to inform the selection process for subsequent calls to the Scheme, in order to build a better understanding of what kind of applications might be successful. This ‘learning curve’ appears at several points in the scheme, so we discuss it subsequently in a separate section.

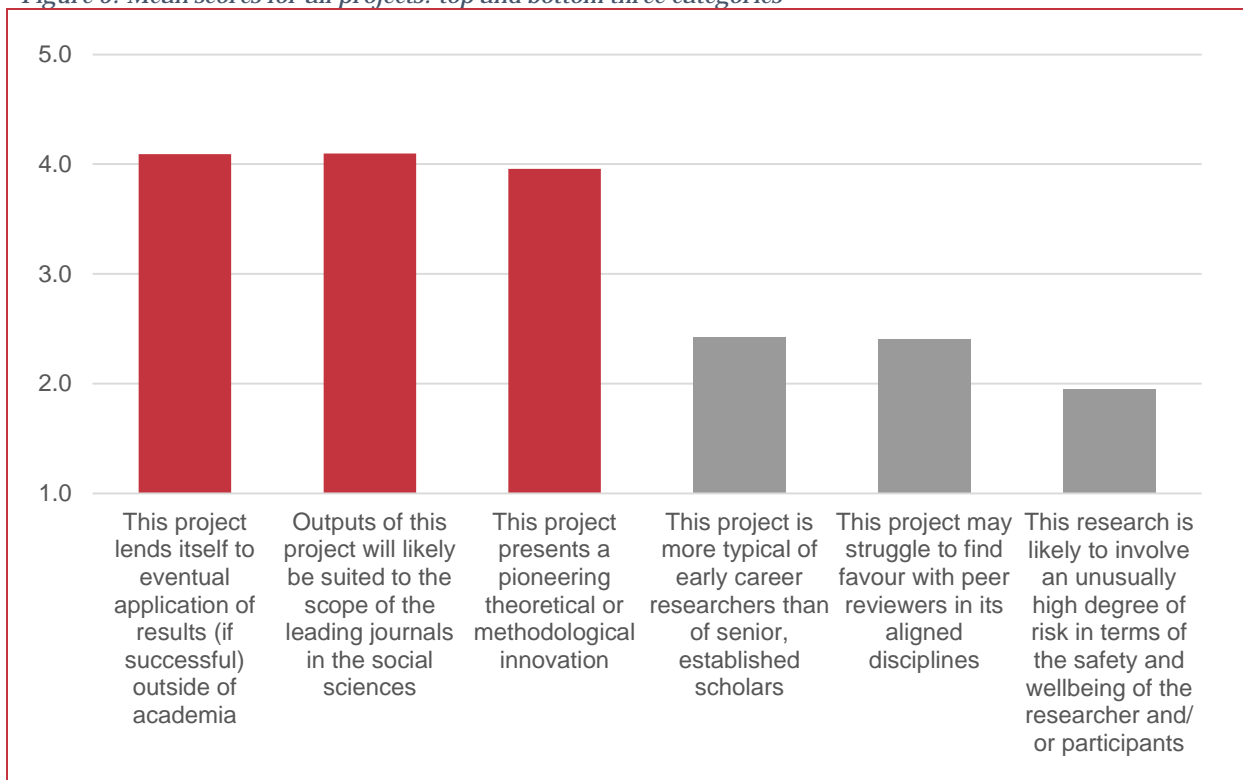
3.1.1 Proposals submitted to the scheme: overall characteristics

“I enjoyed the exercise very much - it makes a change to get to see proposals which veer towards the Completely Bonkers (to use the technical term) when most research proposals today are so very conservative.” (Anonymous reviewer reflecting on our project review exercise)

Following selection through a range of different and evolving internal selection processes, a total of 242 proposals were submitted to the Scheme over the three calls conducted so far. Data from our project review exercise, in which our reviewers scored 93 of these proposals (45 funded, 48 rejected either at sift panel or Pitch-to-peers stage) on a range of different criteria gives us an overall profile of proposals submitted to the scheme.

Figure 6 shows the criteria that received the overall highest and lowest average scores (out of 5) by our reviewers. There is a strong sense that the scheme attracts proposals that appear suited to eventual publication in leading academic journals as well as application outside of academia, whilst also presenting pioneering methodological or theoretical approaches, in line with the ESRC’s definition of transformative research. There is little sense that projects are perceived to be typical of early career researchers, or that their subject matter might struggle with peer reviewers. Moreover, the data suggest that most proposals submitted are ethically sound.

Figure 6: Mean scores for all projects: top and bottom three categories



5= strongly agree; 1= strongly disagree

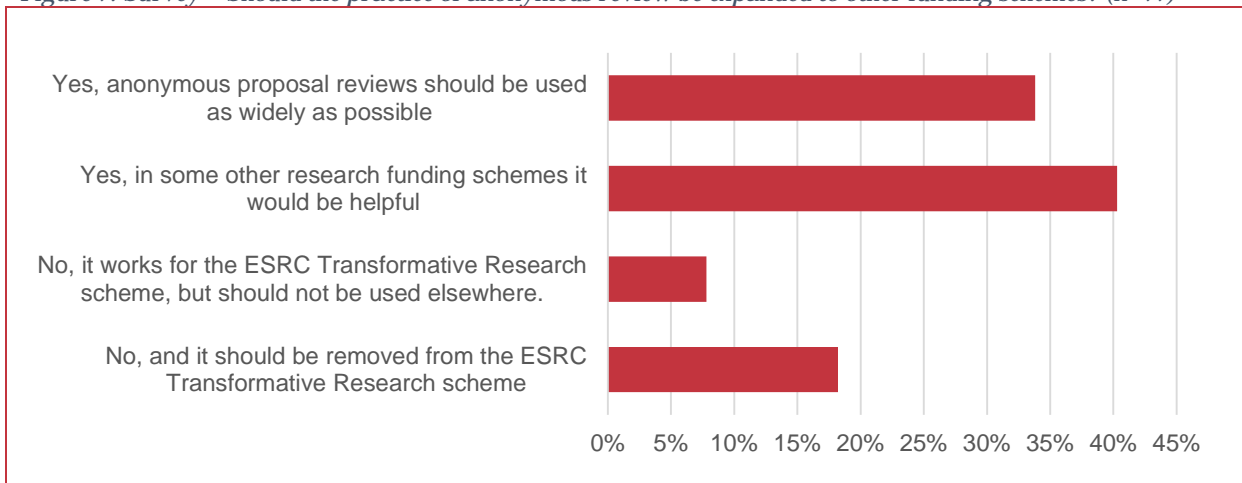
Overall, these results suggest that transformative research in the classical, Kuhnian sense is not noticeably prevalent: rejection from the scientific establishment (journals, peer reviewers) is not characteristic of proposals submitted to the Scheme. Instead, the pool of applications broadly presents ideas that fit into the established professional social science landscape, whilst also containing a strong degree of novelty at the same time.

The Scheme is therefore broadly successful in attracting applications suited to the ESRC’s aims, both in terms of novelty, but also in terms of suiting the ESRC’s remit. The next evident question is whether, from this overall pool of applications, the Scheme’s assessment process adds value by identifying the most transformative ones.

3.2 Stage 2: Sift panel

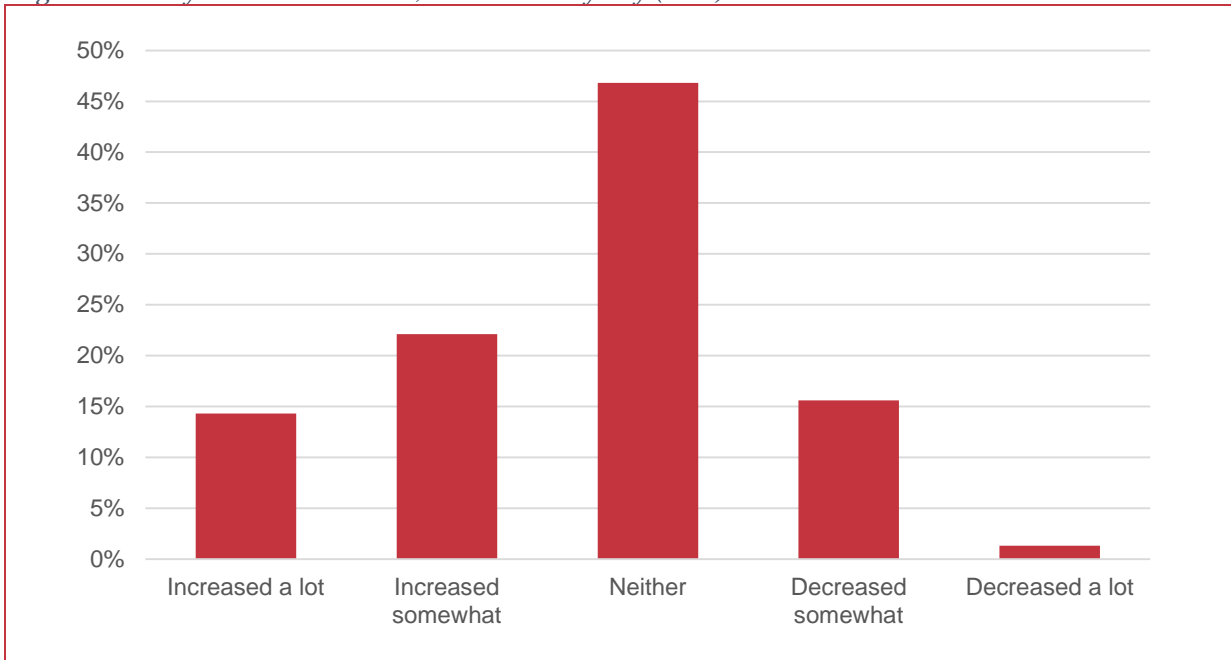
The anonymity of the sift panel review stage of assessment is widely welcomed by several different stakeholder groups. Interviewees generally had a positive view of this step, and our survey of applicants shows overwhelming support for expanding this method at least to some extent.

Figure 7: Survey – Should the practice of anonymous review be expanded to other funding schemes? (n=77)



There is a strong feeling among survey respondents (successful and unsuccessful applicants to the Scheme) that the anonymity factor increased their chances of success, with few having the opposite belief. Comment boxes in the survey, as well as responses from interviewees cited a range of different reasons why this might be the case, ranging from recent maternity leave (meaning lack of recent publication track) to researchers wishing to change direction to a new topic on which they so far had limited established authority. These additional qualitative findings demonstrate that even among established, senior researchers, there can be many possible grounds why assessment procedures emphasising personal credentials can be punitive, and the Scheme appears to have devised a helpful way around these problems.

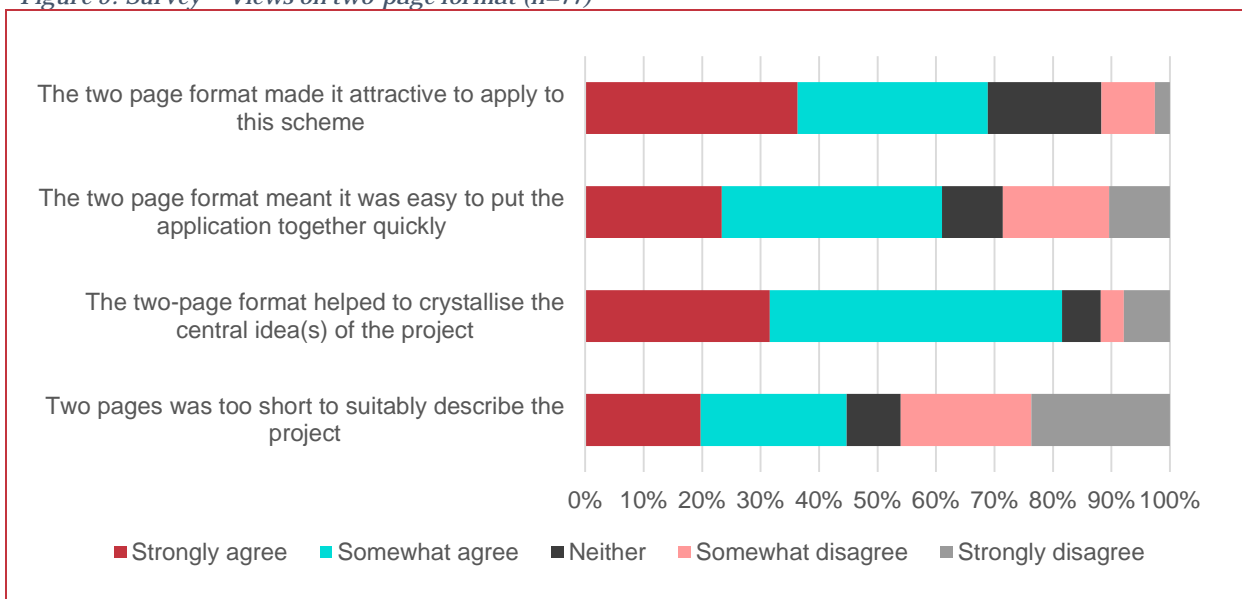
Figure 8: Survey – Chance of success; effects of anonymity (n=77)



No risks or negative points about this anonymous approach are yielded from the point of view of applicants. However, at a higher level it should be noted that this approach does not allow for much control of risk in terms of feasibility and the candidate’s suitability to conduct the research. Though it addresses some important challenges in the contemporary professional academic landscape, this anonymity component needs to be combined with other, more person-centred assessment steps to form a coherent and effective funding tool.

Scheme applicants likewise overwhelmingly viewed the two-page format for applications positively. Despite frequent acknowledgement that such brevity can be challenging, it is generally agreed to be a useful way clearly and succinctly to crystallise research ideas, and also makes it attractive to apply to the Scheme. While challenging, the application process does not appear to be burdensome as such.

Figure 9: Survey – Views on two-page format (n=77)



3.2.1 The sift panel of the 2014-15 call

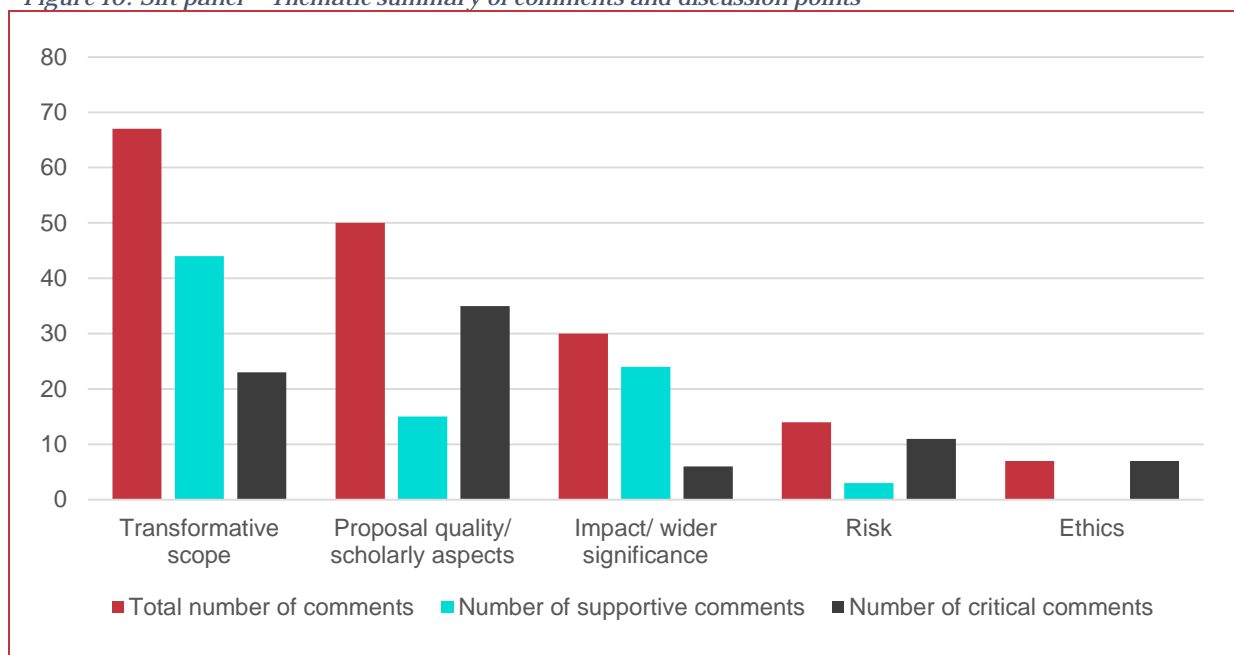
As part of the evaluation, we observed the sift panel event for the 2014-15 call of the Scheme, held at the BIS conference centre in London on 22/04/2015. Besides debate and disagreement on individual proposals, the event was also characterised by a high degree of reflection (including self-reflection and wider reflection on the Scheme) on issues such as application of the ten-point scoring scale, how to define and identify transformative research, the possible meaning of disagreements between assessors, as well as, where applicable, experiences of past calls to the Scheme. These observations are detailed in Appendix E.

During the reviews, it was possible to note the type of comments made for each of the 50 proposals discussed.⁸ The categories chosen (deemed the most suitable during the first few proposal discussions) were:

- Transformative scope of the project;
- Scientific quality and related scholarly aspects of proposal;
- Wider impact and societal importance of the proposed research;
- Risk in terms of feasibility;
- Ethical considerations.

For each assessor, main comments were classified into these categories, noting whether their comments were supportive or critical.

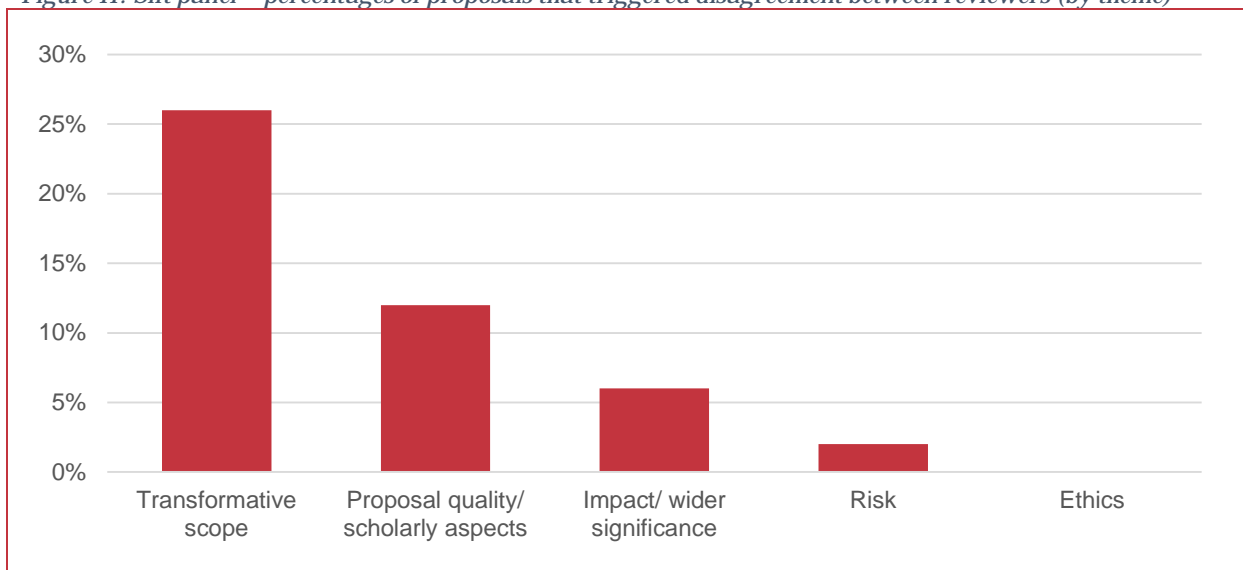
Figure 10: Sift panel – Thematic summary of comments and discussion points



The full data on these observations can be found in Appendix E.

⁸ Though a total of 106 proposals were submitted to the call, not all were discussed: those that had received especially low or especially high scores by all reviewers were categorized accordingly at the start with no further need for discussion.

Figure 11: Sift panel – percentages of proposals that triggered disagreement between reviewers (by theme)



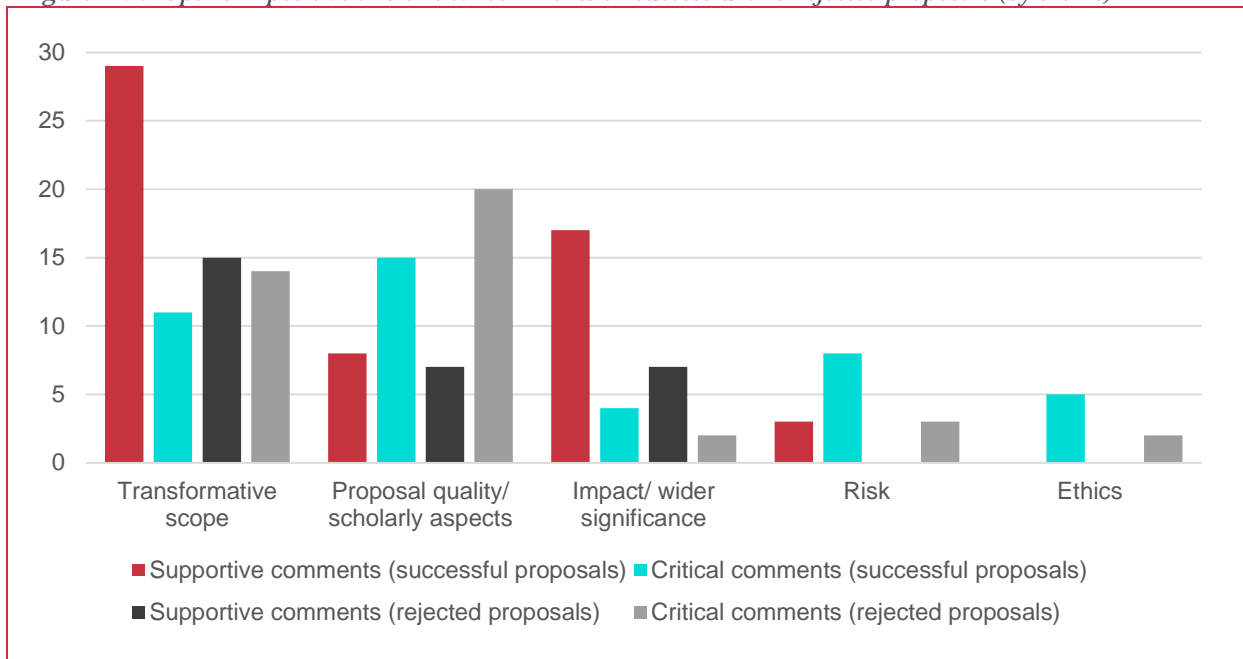
The full data on these observations can be found in Appendix E.

These data are of course not perfect: they numerically summarise often complex arguments, so a degree of interpretation is involved. Nevertheless, the themes and overall character of comments could be determined relatively un-problematically, and the resulting picture highlights some key aspects of the discussion that are also evident from a more qualitative and observational point of view:

- Comments were predominantly focused on the transformative character of proposals;
- Assessors generally pointed to transformative scope as supportive points for shortlisting, but also often problematized the transformative scope of proposals;
- Transformative scope was also by far the most common point of disagreement, with over a quarter of proposals triggering some form of disagreement around whether or not a proposal had genuinely transformative scope. This reflects the fact that the meaning and possible definitions of transformative research was subject to debate more broadly;
- The scholarly quality of proposals was most often a source of criticism;
- Potential wider impact and societal relevance was an important discussion point, with many proposals praised on these criteria, and some criticised for not engaging with such aspects enough;
- Risk was generally seen as a problem, and rarely rewarded, though this was only discussed on a small number of proposals.

Contrasting comments on shortlisted and rejected proposals allows for some further observations:

Figure 12: Sift panel – positive and critical comments on successful and rejected proposals (by theme)



The full data on these observations can be found in Appendix E.

- Successful proposals generally received a greater share of positive comments on their transformative scope;
- Scholarly quality was discussed more in relation to proposals that were ultimately rejected;
- With a very small number of exceptions, risk is generally not commented on in a positive way and is hardly associated with success;
- In a small number of cases, ethical concerns were noted. However, this did not necessarily entail rejection, but was instead most often noted as worthwhile probing at the Pitch-to-peers workshop;
- Both successful and rejected proposals received critical feedback in terms of scholarly quality and transformative scope. Proposals that succeeded were discussed more in terms of their transformative character while unsuccessful ones received more attention to quality.
- Positive remarks about wider impact do not appear to be a good predictor of success at this stage of the scheme.

Overall, data from our project review exercise could not fully corroborate these observations. The review exercise also covers proposals discussed at previous calls' sift panel events, whilst analysis purely for projects at the 2014-15 event would have reduced our sample size to the point where statistical analysis would have become impossible. Our observations alone strongly suggest that the 2014-15 sift panel event managed to strike an appropriate balance between considerations of scholarly quality and efforts to discuss and identify transformative potential. Ethical considerations were also identified, though risk in terms of feasibility did not feature heavily as a discussion point, and was generally not prioritised.

3.3 Stage 3: Pitch-to-peers

"I'm certainly not the same person I was this morning" (Anonymous, applicant after the Pitch-to-peers presentations had finished)

Whilst there is broad consensus about the merits of the anonymous panel review of 2-page applications at the first assessment stage, opinions divide significantly on the second stage, the Pitch-to-peers presentation events. Panellists and ESRC officers generally spoke favourably about the events.

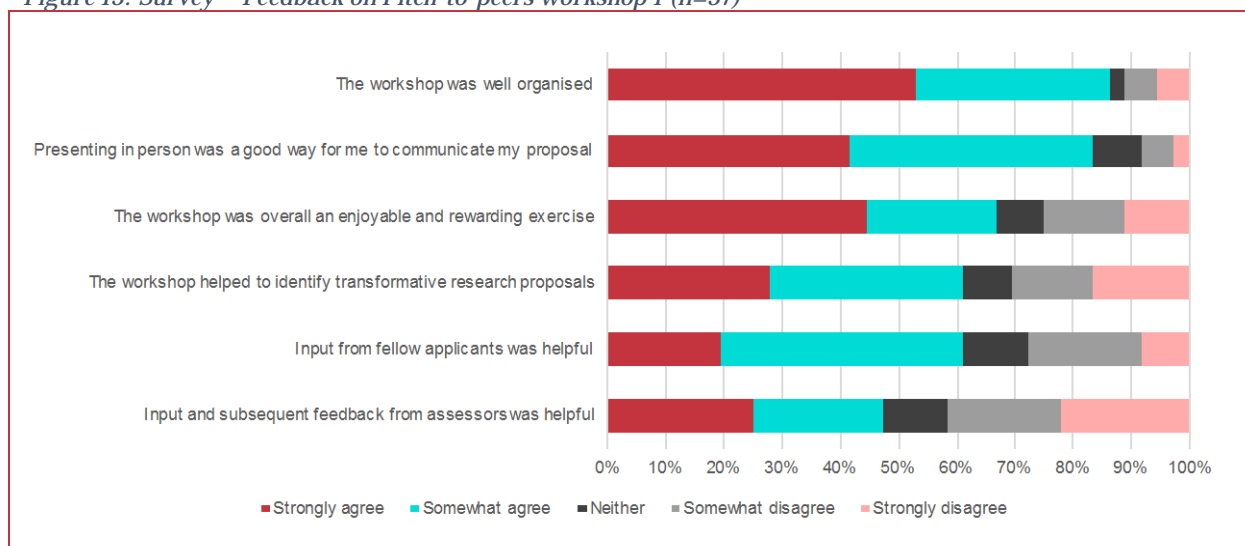
An often-noted benefit was that, following the anonymous panel review, it allowed panellists to meet the applicant in order to get a sense of whether the individual is in fact suitably knowledgeable and capable of carrying out the project. Additionally, some panellists noted that the novelty of proposals to the Scheme tend inevitably to leave many open questions, which are best addressed in an interactive setting.

However, our survey and interviews also generated many comments, about the following areas.

- For individuals suffering from anxiety and related conditions, the workshops can be an intensely stressful experience;
- Given the need to travel and stay overnight, childcare and other duties can become a major problem;
- There is a perceived danger that assessors and/or fellow applicants reward presentation skills rather than substantive ideas;
- The events can become somewhat of an archetypal masculine occasion, with men dominating discussions and generally more socialised into creating a confrontational culture of selling and persuading;
- The events have led to some criticisms in the social science community: overall the format is viewed by some as analogous to a television programme ('Dragon's Den').

However, these points were most often made by individuals who had experienced the first ever Pitch-to-peers workshop (2013), and less often in relation to the second and third (2014 and 2015), indicating once again that over time these events have been improved (see Section 4). Though our survey results show that the majority of attendees are broadly favourable towards the workshops, there are sizeable minorities who are less favourably inclined, and the points above show that these events have in the past entailed serious problems. Whilst these have all but disappeared in more recent calls, maintaining this positive trend is dependent on continuing the current self-reflexive approach of organisers and panellists, and learning the lessons from past failings.

Figure 13: Survey – Feedback on Pitch-to-peers workshop 1 (n=37)



Observations of the Pitch-to-peers event for the 2014-15 call of the Scheme are detailed in Appendix F. Three main points deriving from these observations are worth noting here:

- There need to be clearer 'rules of engagement' – there were a few differences of practice in the three presentation groups but some level of clarification on paper would be helpful. More

generally, issues such as conflicts of interest and the role of fellow applicants' scores were not clear to everyone from the outset;

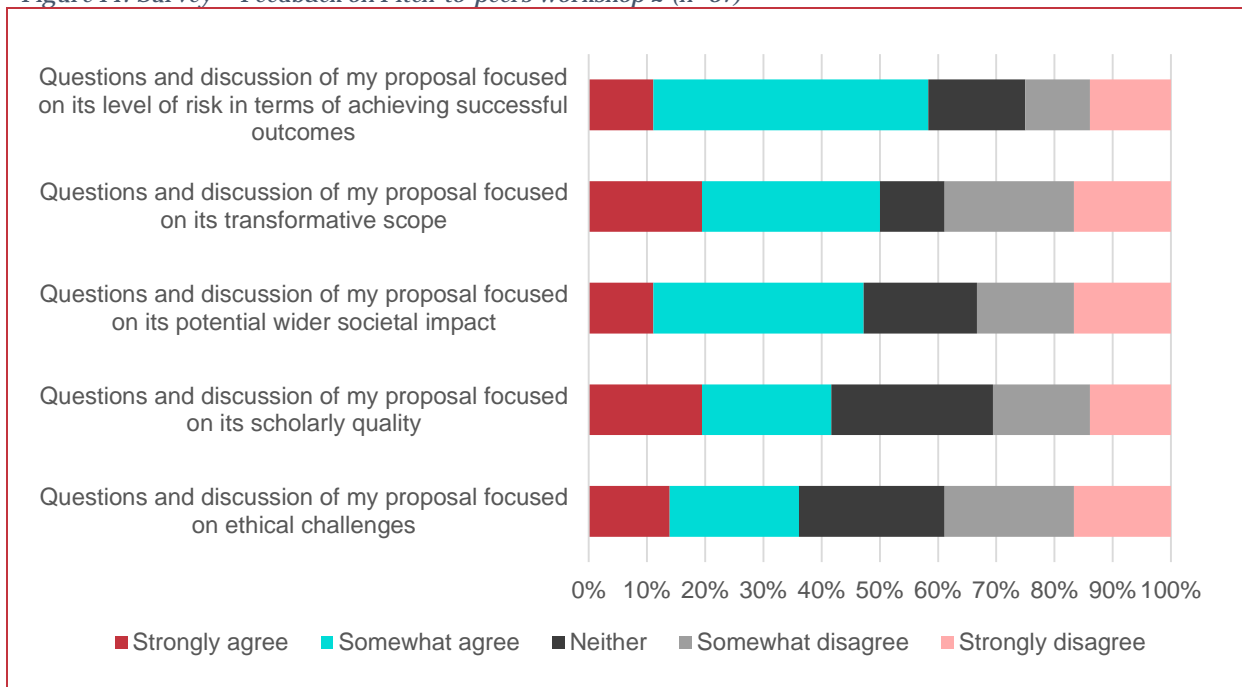
- Observation of the event highlighted that the selection of panellists and especially the chair is critical: a successful event is dependent on the chair's ability to explain the scheme to everyone, and to ensure a relaxed, non-competitive atmosphere. The sub-chairs needed to be selected based on their ability and willingness to take on similar roles, and to encourage collegial discussion rather than 'attack-mode' questioning. This was broadly accomplished here, but it seemed to be dependent on the selection of people, rather than codified process, so it could easily fail on other occasions, depending on people present;
- Throughout, there was an sense of self-reflexivity of the scheme and panel, in the shape of self-scrutiny and discussion on how best to behave and approach the event, sometimes reflecting on past Pitch-to-peers events. The Chair was conscious of her role to ensure the event worked, to check for feedback, suggestions for improvement, and prior concerns of applicants and panellists.

3.3.1 The function of Pitch-to-peers

Whilst the sift panel stage clearly had a strong emphasis on identifying and problematizing the transformative scope of proposals, the rationale of meeting the applicant in person and checking their suitability of the task and asking questions left open by the written application suggests different aims at the Pitch-to-peers stage. The EPSRC's IDEAS Scheme, which has a similar set-up to the ESRC's Scheme and partially acted as a model, has venture capitalists represented on the assessment panel for presentations. Having identified proposals with transformative scope at the sift panel, the presentation event is therefore intended to control for risk.

To an extent this is also true with the ESRC's Scheme, though not to the same extent. Our survey results do show that risk and feasibility were a prominent theme of questions asked.

Figure 14: Survey – Feedback on Pitch-to-peers workshop 2 (n=37)



However, these findings are not fully corroborated by our observation of the 2014-15 Pitch-to-peers event. We coded each question in response to each presentation we attended (18 out of the 25) into the following categories:

- Was the question asked by a fellow applicant or by an assessment panel member?

- Did the question invite an expository or a defensive response?⁹
- Was the question related to transformative scope, scholarly quality, non-academic impact, risk in terms of feasibility, or ethics and safety?

More so than our survey data, the observations show that transformative scope of proposals was not the dominant theme of questions asked. However, rather than risk, it was scholarly quality that formed the main focus of questions.

Table 9: Profile of questions asked at Pitch-to-peers event

Presentation details		Question details								
Presentation categories	Number of Questions asked (mean)	Question from		Character of question		Theme of question				
		Panellist	Applicant	Defensive	Expository	Transformative scope	Scholarly quality	Wider Impact	Risk/ Feasibility	Ethics
All	8.5	68%	28%	30%	70%	17%	53%	6%	12%	12%
Observer 1*	7.8	69%	31%	26%	74%	18%	55%	5%	11%	12%
Observer 2	9.1	67%	25%	34%	66%	16%	52%	7%	14%	12%
Accepted	8.4	63%	37%	21%	79%	11%	60%	3%	16%	11%
Rejected	8.6	72%	20%	39%	61%	22%	47%	9%	9%	12%
Male appl.	8.5	68%	28%	33%	67%	24%	50%	7%	10%	10%
Female appl.	8.4	68%	28%	28%	72%	11%	57%	5%	14%	13%
Rover 1	8.0	68%	32%	22%	78%	10%	50%	11%	12%	17%
Rover 2	6.7	60%	40%	30%	70%	32%	49%	0%	10%	10%
Rover 3	9.5	71%	20%	37%	63%	17%	57%	5%	14%	8%
Group 1	8.3	77%	23%	20%	80%	12%	50%	13%	11%	13%
Group 2	8.8	71%	29%	35%	65%	9%	62%	6%	13%	12%
Group 3	8.2	59%	29%	30%	70%	30%	44%	3%	12%	10%

*As there were three different presentation groups, based in separate rooms, the two present observers from Technopolis scored different presentations. Their respective scores are noted here to indicate that there is little cause to suppose that the scoring system was used differently.

These data of course only relate to the 2014-15 event whilst the survey data cover the two previous ones. Moreover, as with our analysis of the sift panel meeting, a degree of interpretation is inherent in the observation data. Nevertheless, both survey and observation data suggest that unlike at the sift panel, transformative scope is not the primary focus at the Pitch-to-peers events. Instead, these events

⁹ To distinguish between these, we understood 'expository' as questions that focused on the applicant's presentation and asked to further explain an aspect of it. 'Defensive' denoted questions that brought in ideas or concepts not contained in the applicant's presentation itself; e.g., 'could you please explain more/ tell us why you are doing x?' is expository, 'author x/ theory y says that there is a problem with your plan. What do you say to that?' is defensive.

are best described as ensuring robustness of proposals, by prioritising issues around feasibility or scholarly quality and ability of applicants.

A few further observations are worth noting based on our observations:

- There were significant differences in the number of questions asked for each presentation depending on which rover panel of assessors was present. This reflects the more general problem that there was little sense of a standard format across the event;
- Fellow applicants asked 28% of observed questions. However, this also differed significantly between rover panels and groups;¹⁰
- Applicants who were ultimately successful generally had a higher proportion of questions from fellow applicants;
- Gender of the presenter did not appear to play a part in terms of the kind of questions that were asked, though questions about transformative scope of the proposed research were levelled at female applicants slightly less often.

Inherently, there is some concern – highlighted also by our interviews – that the competitive element of the Pitch-to-peers events might lead to fellow applicants being especially challenging towards each other. Our observation data do not confirm this. In fact, fellow applicants were far less likely to ask questions inviting defensive responses. These findings underscore that a somewhat collegial atmosphere had been facilitated.

Table 10: Character of questions – applicants and panellists compared

	Questions from fellow applicants	Questions from assessment panellists
Questions inviting expository response	83.3%	66.0%
Questions inviting defensive response	16.7%	34.0%
	100%	100%

¹⁰ 'Rover panels' are groups of assessors who move between the three rooms in which presentations take place; 'groups' refers to the three rooms themselves, with each containing a fixed group of presenters and an assessment panel member in charge of that particular room.

4 The three calls compared: understanding the Scheme as a learning process

Throughout the evaluation it became increasingly apparent that there has been considerable development and continuous learning among stakeholders in the Scheme. That is to say, the Scheme began with relatively broadly set aims and goals, with key terms loosely defined, notably ‘transformative research’ itself, assessment procedures un-tested, and applicants and institutions unsure about what projects to put forward and how to increase chances of success. Over time, stakeholders have learned, concepts have become better understood and problems have been addressed. Consistently, our analysis has highlighted trends and differences across the three calls conducted so far, most often pointing into a promising direction. Few explicit changes have been made to the Scheme over the three calls. Rather than incremental substantive modifications to the ‘written rules’, we observe instead increasing clarification of what may be termed the Scheme’s ‘un-written rules’, leading to more positive outcomes. We present key findings on this matter in the sub-sections below.

4.1 Ensuring a level playing field

ESRC’s own data suggest that on key criteria the pilot call in 2012/13 appeared to favour men and senior researchers at both main assessment stages; however there was little evidence of disadvantage either for women or for early career researchers by the 2014-15 call. Indeed, even in terms of the profile of applicants (“Total submitted”) a degree of equalisation took place over the three calls.

Interviews did not highlight any particular steps taken in terms of ensuring a more diverse range of applicants, though the dangers around the Pitch-to-peers events descending into overly aggressive or confrontational occasions was acknowledged to have been noted after the first call, and attention was paid to mitigating this danger subsequently.

Figure 15: Success rates of the 2012/13 call by gender and seniority

	Submitted	Shortlisted	Funded
Total	67 (100%)	32 (100%)	20 (100%)
Female	38%	32%	32%
<i>Female, Early career</i>	5%	3%	5%
<i>Female, Mid-career</i>	12%	6%	0%
<i>Female, Advanced career</i>	21%	23%	26%
Male	62%	68%	68%
<i>Male, Early career</i>	5%	3%	5%
<i>Male, Mid-career</i>	20%	23%	16%
<i>Male, Advanced career</i>	38%	42%	47%
Early career	9%	6%	11%
Mid-career (Lecturer or above)	32%	29%	16%
Advanced career (Professor)	59%	65%	74%

Source: ESRC data

Figure 16: Success rates of the 2013/14 call by gender and seniority

	Submitted	Shortlisted	Funded
Total	69 (100%)	26 (100%)	13 (100%)
Female	44%	42%	46%
<i>Female, Early career</i>	7%	8%	8%
<i>Female, Mid-career</i>	13%	12%	23%
<i>Female, Advanced career</i>	24%	23%	15%
Male	56%	58%	54%
<i>Male, Early career</i>	13%	23%	38%
<i>Male, Mid-career</i>	19%	23%	8%
<i>Male, Advanced career</i>	24%	12%	8%
Early career	21%	31%	46%
Mid-career (Lecturer or above)	32%	35%	31%
Advanced career (Professor)	47%	35%	23%

Source: ESRC data

Table 11: Success rates of the 2014-15 call by gender and seniority

	Submitted	Shortlisted	Funded
Total	106 (100%)	25 (100%)	12 (100%)
Female	46%	52%	58%
<i>Female, Early career</i>	12%	8%	17%
<i>Female, Mid-career</i>	18%	24%	17%
<i>Female, Advanced career</i>	16%	20%	25%
Male	54%	48%	42%
<i>Male, Early career</i>	11%	16%	17%
<i>Male, Mid-career</i>	22%	16%	17%
<i>Male, Advanced career</i>	21%	16%	8%
Early career	22%	24%	33%
Mid-career (Lecturer or above)	41%	40%	33%
Advanced career (Professor)	37%	36%	33%

Source: ESRC data

4.2 Rising consensus between panellists and fellow applicants at Pitch-to-peers events

Proposal presentations at the Pitch-to-peers events have an audience of fellow applicants as well as assessment panellists (the same who conduct the first assessment phase). All are invited to score each proposal, though panel members subsequently decide which proposals to accept, and are under no obligation to take scores by other applicants into consideration (though they can do so at will). A brief

comparison of panel and fellow applicant (peer) scores for each of the three Pitch-to-peers workshops to date show that this is significant. Especially in the first event, fellow applicants judged very differently from the panel, meaning that more systematic inclusion of their scores might well have had considerable influence on the outcomes (see Appendix G). Fellow applicants also generally gave lower scores than panel members, resulting in a lower number of proposals that would have reached the pass mark of 7. However, both in terms of overall average scores, and in terms of decisions on individual projects, divergences between panel and peers have become less dramatic over time. Most recently, fellow applicants in fact scored overall more favourably than panellists.

Table 12: Comparing Panel and fellow applicant mean scores (out of 10)

	Panel mean	Peer mean	Total funded	Funded (peers)*
2012-13 call	6.37	5.798	20	7
2013-14 call	6.59	6.61	13	9
2014-15 call	6.348	6.868	12	12

*Hypothetically – assuming peers’ judgement alone with a mean pass-mark of 7 were the sole criterion

Not only did fellow applicants’ scores become less punitive compared with panellists’ scores; over the three calls, there is evidence of increasing consensus between peers and panellists at the Pitch-to-peers events on a project-by-project basis. In the earliest call there is no significant correlation between peer and panel means, in the second there is a moderate correlation, and in the most recent call this correlation becomes strong.

Figure 17: Correlation between Peer and Panel scores at Pitch-to-peers events

Call year	r-value	Correlation
2012-13	0.30	None (weak)
2013-14	0.52	Moderate
2014-15	0.63	Strong

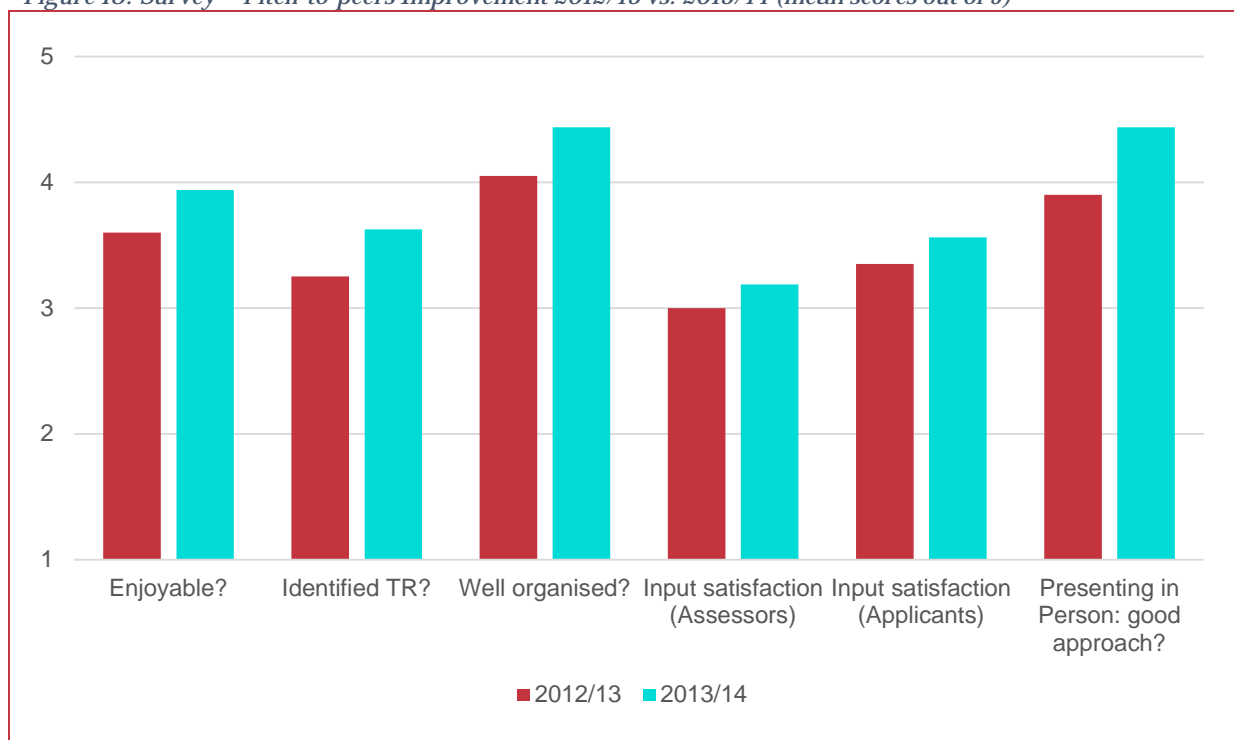
Data generated as part of the project review exercise

These findings reflect once again the sense of learning throughout this Scheme: this level of convergence suggests the development of greater understanding over time, where at some level the various groups involved in the Scheme become better able to reach consensus. A tentative hypothesis here might be that a certain culture of transformative research has arisen around the Scheme (so among applicants and panellists especially), which is slowly gaining traction.

4.2.1 Feedback on Pitch-to-peers

Aside from the outcomes of the Pitch-to-peers workshops, our survey results are also suggestive of overall increasing satisfaction with the events between the first and second call. We asked former applicants to score out of five the extent to which they found it enjoyable, whether it succeeded in identifying transformative research, whether it was well organised, whether feedback and questions from panellists and fellow applicants respectively were satisfactory, and more broadly, whether presenting in person was felt to be a good approach. There is a potential limitation with regard to sampling and reliability with the results, and statistical tests are not possible with such a small sample and population, but the results nevertheless support the idea of improvement over time on all criteria.

Figure 18: Survey – Pitch-to-peers Improvement 2012/13 vs. 2013/14 (mean scores out of 5)



5 denotes 'Strongly agree', 1 denotes 'strongly disagree', making 3 the neutral line.

Additionally, we found some illuminating data on the relationship between applicant’s place in the day’s running order of presentations and their outcome. At all three Pitch-to-peers workshops to date, applicants were in a group of between eight and eleven individuals due to present. Concerningly, we find that in the first ever Pitch-to-peers workshop, applicants who presented towards the end of the running order, i.e. later in the day, were far less successful. However, in subsequent events there is far less evidence for this. A key explanation for this might well lie in the comments from our interviews and survey, which noted a more stressful atmosphere at the 2012-13 event. However, given the overall higher number of projects funded at the first event (20 in 2012-13, as opposed to 13 and 12 in the subsequent two calls), it is likewise possible that assessors became less lenient as the day progressed due to emerging budgetary constraints. In either case, unfair advantage based on an applicant’s place in the day’s running order virtually disappears after the first call.

Table 13: Disadvantage of presenting late in the day

Pitch-to-peers event	Funded projects: mean position in the day’s running order	Unsuccessful projects: mean position in the day’s running order	Difference
2012-13 call	4.80	7.58	2.78
2013-14 call	4.77	5.08	0.31
2014-15 call	4.25	5.08	0.83

4.3 Other learning curves and feedback loops

In addition to the findings above, the idea of the Scheme as a learning process became evident at several other points, some of which were alluded to earlier. Interviews and observation also highlighted:

- Many institutions involve former applicants in their selection of proposals to put forward for the Scheme, in order to better understand what the Scheme is looking for and which applications might have the best chances of success;
- Similarly, several grant winners of past calls have subsequently been asked by the ESRC to act as panellists for later calls. At the sift panel and Pitch-to-peers meetings, these individuals were invited to talk about their experience and issues highlighted by their experience;
- Following and concerns about creating an excessively high-pressure atmosphere after the first ever Pitch-to-peers event, panellists are aware of the need to ensure a suitably open and relaxed atmosphere. This is especially true of the most recent call's panel chair, who noted on several occasions the importance of managing the event properly and keeping any sense of competition between applicants to an absolute minimum;
- Institutional eligibility to the Scheme was strongly prohibitive for the first call: only institutions that already received significant ESRC funding through other channels were eligible to submit. In the absence of clear rationales for this, along with the realisation that successful applications did not necessarily come from overall high-performing institutions, these eligibility rules have been relaxed from call to call, as far as the relatively small budget for the Scheme permits.

4.3.1 Reflexivity of the assessment panel

"As soon as we unpack something, we begin to talk it down" (Anonymous, Panellist reflecting to others at the sift panel)

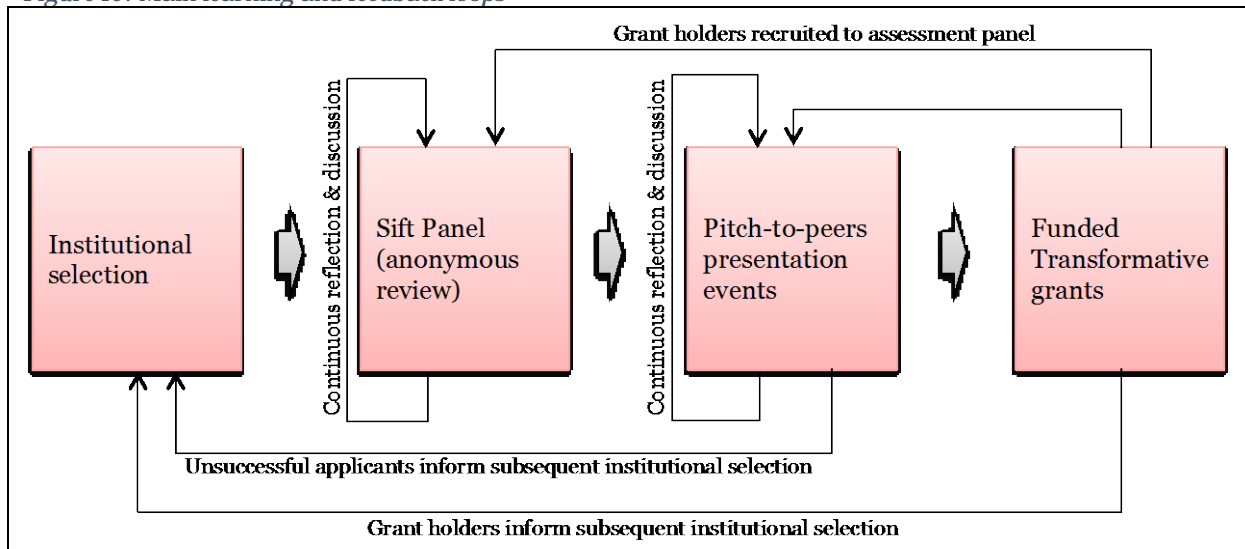
Finally, it is worth noting that we observed a high level of learning and self-reflection at the micro-level. As the observation data for the sift panel (Section 3.2.1) and the Pitch-to-peers event (Appendix F) show, there were many inclusive discussions, where panellists and chairs raised concerns, scrutinised procedures and definitions, as well as their own behaviour. A few pertinent examples highlighting this sense of self-reflexivity are worth noting here:

- It was noted on two occasions that mean scores are a problematic point of reference: transformative research is likely to divide opinion in some form. The chair reaffirmed this by noting that divergence between assessors' scores should be taken seriously whenever spotted;
- Panellists noted that it was difficult to strike a balance between risk and credibility;
- The problematic distinction between 'new/novel' and 'transformative' was noted on a few occasions: all funded research ought to add something new in some form – so where is the cut-off line?
- It was fairly clear that the 10-point scale was not used in the same way by all panellists, with some openly admitting they only marked as high as 7 as a rule, and others noting they gave the highest possible marks to encourage transformative ideas;
- At the Pitch-to-peers event, the question was raised of what fellow applicant's scores should be for. After some discussion, it was agreed that these should be used in cases where panel scores disagree, to help come to a decision;
- Following the presentations, the Chair held a de-briefing with all applicants and panellists. This yielded a few points:
 - Assessment criteria (the 10-point scale) were noted as being too fine-grained. 10 points were too much to reliably classify proposals;

- Shorter scales, but for several different criteria were mooted as an alternative, so e.g. a 3-5 point scale for TR / quality / other aspects;
- Adding a confidence level to judgements was also noted as a possibility, as attendees felt they could judge some presentations much better than others.

Though our findings provide evidence for a wide and varied array of learning processes over time, the main demonstrable learning and feedback loops can be schematised and are shown in Figure 19.

Figure 19: Main learning and feedback loops



5 Perception and attractiveness of the Scheme

5.1 Individuals

Though the first call of the Scheme attracted mainly senior researchers from predominantly high-ranking institutions, and was also heavily dominated by male applicants, the Scheme has since broadened its intake towards greater representation, in terms of seniority, gender and level of host institutions. Our interviews with university research officers and grant holders alike suggest that the Scheme is not understood to be reserved for any particular group or ‘type’ of researcher.

In terms of the Scheme’s appeal and attractiveness to individual researchers, our survey and interview data show an overall positive picture. However, there are also some areas of concern. The final section of our survey of applicants invited further comments on the Scheme. Many of these were positive, praising the various aspects of the Scheme noted in this report. However, some were also more cautious or even critical. Two points were particularly evident: first, some applicants felt that feedback from reviewers had been either inadequate or difficult to access; second, the resource constraints were problematised, especially in the context of an apparent lack of follow-up funding.

Figure 20: Survey – Further comments on the Scheme: Themes and frequency

Theme	Frequency
Positive	13
Cautious (mixed positive/critical)	5
Critical	7

Especially the two-page format for applications, as well as the anonymous review stage are generally viewed favourably, as these two features respectively reduce the amount of time needed to produce applications, and allow researchers who for a broad range of reasons might not have recent track in the topic they are proposing to have a chance of success (these advantages were already noted in Section 3).

Some interviewees commented that ESRC transformative grants have a certain level of prestige in the social scientific community, whilst others found that the Scheme is too low-profile so far to play a particular role in a researchers’ standing among peers. The Scheme does not appear to be related to particular career stages but was generally agreed to be as useful as other grants of comparable size to career development. Critically in this respect, a few grant holders noted that the anonymity of the first stage allowed them to change from their established expertise to a different subject matter without being punished for their lack of track. In this sense, the Scheme provides a tool for greater flexibility in researchers’ careers, which can otherwise easily become path-dependent.

5.1.1 Lack of wider context and integration

The size and duration of the grants was a matter of some dispute: whilst some grant holders and panellists found these parameters too restrictive, others felt it to be adequate in order to produce a base of evidence suitable to verify the transformative idea. At the heart of this debate lies the current lack of immediately obvious follow-up funding. The perception of many interviewees and survey respondents is that this Scheme has a stand-alone character, placing a burden on the researcher to achieve ambitious results with limited resources. The relatively modest grants are necessary due to overall funding budget constraints, and also in order to limit the extent of risk associated with transformative research. There may be a need to contemplate clearer pathways to follow-up funding, e.g. for cases where funded projects show promising outcomes but risk being stifled due to lack of resources. Various interviewees mooted the possibility of allowing transformative grant holders to be fast-tracked through other ESRC schemes under certain conditions. This integration of the Scheme into a broader funding landscape is therefore worthwhile considering in possible future calls. This is

further underscored by our project review exercise, in which the potential need for follow-up funding for successful results was identified as one of the key features associated with successful applications (see Section 2.1).

5.2 Institutions

From the institutional level, our interviews highlighted a similarly favourable view of the Scheme. In particular, it was felt that the Scheme offers a helpful mitigation to current trends of concentration of funding into fewer, larger grants, resulting in a 'raised floor'. The smaller size of transformative grants is therefore helpful in this respect. Most institutional officers we spoke to very much want the Scheme to remain in place, and the loosening over time of institutional eligibility was widely welcomed.

A problem from the institutional perspective appears to be lack of clear definition of what the Scheme is looking for. There is an evident perception that it stands in contrast to response mode grants, but beyond this, institutions struggle to understand what proposals have the highest chance of succeeding. As highlighted in Section 4, the continuous learning and feedback occurring around the Scheme is likely to remedy this in the future. However, interviewees suggested ways of speeding up this process: several mentioned that a clear and detailed definition of transformative research, or alternatively circulation of information about successful grants would help the internal selection process, as this could facilitate better narrowing down of potential applicants.

5.2.1 *The 2012-13 institutional payments*

In the first call, grant-winning institutions received an additional institutional payment of £50,000, besides the project grant itself.

Grant holders are not well connected to these sums. According to our survey data, 23.5% of grant holders from the 2012-13 call who responded to our survey were not aware that this additional payment even existed. Only a handful (out of a maximum possible 20) of 2012-13 grant holders chose to share their knowledge about what the additional sums were used for. Answers include:

1. The institutional payment was used to support the development of transformative research projects and to fund a series of sandpit events.
2. The institutional payment 'disappeared' into the School. The grant holder subsequently heard that it had financed a new teaching module.
3. The institutional payment was divided into smaller pots in a Transformative Social Science Scheme for seedcorn, networking, sandpit and development activities.
4. The institutional payment was used to support transformative research in the University. Staff were allowed to bid for funds, for example, to re-skill in interdisciplinary methods, or to pilot a transformative study. The grant holder put in a bid for some additional monies, which were awarded to slightly extend the original project.

Our interviews identified especially notable use at the University of Aberystwyth, where the money was used in part for a symposium on transformative research, and also to part-fund a position for Professor of Transformative Social Science.¹¹

¹¹ see <http://www.aber.ac.uk/en/news/archive/2014/03/title-148301-en.html>

Figure 21: Aberystwyth University Transformative Research Symposium

The screenshot shows the Aberystwyth University website. The header is yellow with the university logo on the left and a search bar on the right. Below the header is a purple navigation bar with links for 'STUDY WITH US', 'INTERNATIONAL', 'RESEARCH', 'NEWS', 'ALUMNI', and 'ABOUT'. A purple sidebar on the left lists various services under 'Research, Business & Innovation'. The main content area is white and features a 'Call for Proposals - Funding for an International Symposium' section. This section includes a heading, three key questions, a paragraph about the ESRC funding, a paragraph about the symposium's aim, a paragraph about the funding's use, and a paragraph about the bid submission process. Contact information for Michael Woods is provided at the bottom of the main content area.

source: <http://www.aber.ac.uk>

Whilst there is therefore evidence of suitable use of such payments, more transparency would be needed if this practice were to be reinstated. The lack of connection between grant holders and institutional payments suggests that simple attachment to winning grants is not an intuitively suitable way of allocating such payments.

5.3 Success rates

It is worth briefly considering the Scheme's success rate as an additional marker of attractiveness. Whilst the first call had a high overall success rate, this has decreased considerably since then. Current Scheme totals, as well as individual calls, consistently have lower success rates than ESRC response mode grant funding. This suggests, along with the increase in applications over the three calls, that there may well be scope for expanding the Scheme.

Table 14: Scheme success rates and contrast to ESRC totals

	Total applications	Total awards	Success rate
Scheme Totals	242	45	19%
2012-13 call	67	20	30%
2013-14 call	69	13	19%
2014-15 call	106	12	11%

	Total applications	Total awards	Success rate
<i>ESRC Response mode totals 2012-13*</i>	<i>315</i>	<i>84</i>	<i>27%</i>
<i>ESRC Response mode totals 2013-14*</i>	<i>334</i>	<i>82</i>	<i>25%</i>
<i>ESRC Response mode totals 2014-15*</i>	<i>420</i>	<i>53</i>	<i>13%</i>
3-year average ESRC response mode totals	356	73	22%

*ESRC annual reports and accounts

6 Phase 2: Outputs, Outcomes and future perspectives

Phase 2 of our evaluation ran from January to June 2016 and focused on the outputs and outcomes of projects funded under the Transformative Research scheme. Besides a more general assessment of outcomes of the scheme, supplementary issues of interest for this phase are whether any projects funded through the Scheme failed to meet their objectives and if so, why, as well as whether any genuinely transformative outcomes have emerged and how these might be developed further.

The timing of the second phase was set in such a way that projects funded in both the 2012-13 and 2013-14 calls were completed (with an exception of a small number of 2013-14 grants, which had been significantly extended) and could therefore be included in our analysis. Our findings on outcomes and outputs do not include projects funded in the 2014-15 call, as these are still in progress. Our data collection and analysis for Phase 2 consisted of three main components:

- Interviews with 20 grant holders (12 from the first call and eight from the second)
- Analysis of outputs and key findings data entered by grant holders into ResearchFish
- A peer review exercise of 80 outputs (journal articles and working papers): 40 listed on ResearchFish as outputs from Transformative Research grants and 40 listed as outputs from ESRC standard response mode (hereafter: Standard) grants. Reviews were conducted by a panel of six peer reviewers (all of whom had already contributed to our Phase 1 review of applications), who, for each output, completed a template similar to the one used in our Phase 1 review: 15 characteristics pertaining to transformative scope, and an overall judgement on whether or not the output is transformative. Additionally, we added a REF-style quality-profiling matrix, and also asked reviewers to rate their own confidence in their judgement for each output. Reviewers were not told which outputs were from the TR scheme, or even that half of the outputs were not, though some disclosure was inevitable, as some outputs included an acknowledgement of the TR funding scheme.

We summarise our headline conclusions here at the outset:

- In terms of productivity, the TR grants are roughly equal to standard grants, although they generally produce similar quantities of outputs in a shorter time and on a slightly lower average budget.
- Likewise in terms of quality, there is no apparent difference between Standard and TR grant outputs.
- The vast majority of projects were successful at some level, albeit in many cases only as a proof-of-concept. There is little sense of undue risks or absent rewards.
- On several criteria, the TR projects and individual outputs are judged more transformative than their counterparts in the Standard grants scheme, especially in terms of unusual interdisciplinary perspectives and in their capacity to challenge fundamental assumptions of social scientific enquiry and lead to paradigm shifts. The TR grants also produce outputs that more often tend to suggest significant further work or follow-on funding may be necessary for results to come to full fruition.
- However, this does not mean that TR grants categorically produce 'transformative' outputs while standard grants do not: many outputs in both categories were identified as transformative, though the TR grants considerably more so. On the criteria we measured, many outputs of Standard grants also have characteristics associated with being 'transformative'.
- There are strong relationships between the transformative character and scope of the grant outputs, and their quality. Moreover, there is some evidence that outputs associated with early career researchers do less well in both these sets of dimensions. Such tendencies are to be expected, however, as early career stages might denote less experience in rapid production and publication of high-quality outputs.

- Particularly for the more recent grants, results and critical outputs are still emerging. Even for the less recent grants, there are still ‘publication tails’ in progress. Many grant holders from both rounds analysed here have secured further funding from a range of different funders. Most grant holders are moreover either in the process of writing applications for further funding, have applications under review or are considering application in the near future. This mitigates to a large extent any concerns that might otherwise arise around the increased proportion of awards to ECRs in the later calls of the scheme: initial lack of experience to produce high quality outputs is likely to diminish as agendas progress.
- Relatedly, there is an overwhelming sense that funded projects are seen by grant holders as merely marking the beginning of a longer-term research agenda, rather than being discreet projects with fixed end-points.

6.1 Productivity of the grants

Initially we sought to obtain an overview of the productivity of the TR grants and to assess whether levels of outputs are at all comparable to equivalent standard grants. The ResearchFish data were an important resource for this task. Four projects from the 2013-14 round of the TR scheme are still on-going due to extensions, so these were excluded from this part of our analysis, leaving 29 grants. We compare the output levels of these to a comparator selection of ESRC Standard grants: all Standard grants of a value comparable to the TR grants (£180k-£320k) and with a comparable timeframe (start dates between 01-06-2012 and 06-01-2014; end dates between 30-11-2014 and 31-03-2016) formed a pool of 34 ESRC Standard grants as the basis for our comparison. We detail the mean outputs per grant as listed on ResearchFish in Table 15.

Table 15: Mean outputs listed in ResearchFish

	Journal Articles	Conference Abstracts	Book Chapters	Books (incl. edited & monographs)	Working Papers	Reports & other outputs	Total outputs	% of projects with no listed outputs
TR (29)	1.83*	0.93	0.48	0.14	0.03	0.24	3.66	21%
Standard (34)	2.59*	0.97	0.06	0.03	0.15	0.15	3.94	24%

Start dates: Standard: 01/06/2012 – 06/01/2014; TR: 01/06/2013 – 24/11/2014. End dates for both: 30/11 2014 – 31/03 2016

*Mean values are skewed by a small number of projects with high article outputs (typically economics or psychology). Median value for TR and Standard grants is 1

These figures do not suggest that TR grants are in any way different in terms of productivity from their Standard grant counterparts. Especially noteworthy is the fact that the proportions of grants with no listed outputs are also comparable: if lack of outputs signals a failure to achieve results, then these data do not suggest that any undue risks were taken by ESRC through this scheme.

However, we caution that our interviews in conjunction with the above data suggest that not all outputs are in fact logged on ResearchFish. Our interviewees also frequently noted that many outputs – particularly journal articles – are still in progress or under review (an issue we return to later). There is little reason to suppose that these features are any different for Standard grants, so the above data should not be understood in absolute, but only in comparative terms. It is highly likely that the percentages of projects with no outputs listed will drop for both types of grants whilst all other figures will rise further in time, all of which is compounded by the fact that some outputs will not be listed at all, as data entry into ResearchFish does not yet appear to be fully comprehensive on the part of all users.

Beyond mere numbers of outputs, it is also critical to get a sense of quality. ResearchFish invites users to enter journal titles, which could lend itself to an analysis of impact factors. However, particularly in social sciences, there is much critical debate around impact factors as markers of quality, evidenced not least by the San Francisco Declaration on Research Assessment (DORA). Additionally, there are widely acknowledged problems around impact factors or ‘prestige journals’ more generally, and especially when it comes to interdisciplinary endeavours (Rafols et al 2012), which, as we show subsequently, is a significant element in the TR outputs. We therefore opted against a bibliometrics-led analysis of quality.

Instead, we added to our review exercise of outputs a small quality assessment. Our reviewers (all of whom were formerly RAE/REF panellists) were asked to rate each output on a five-point scale on ‘Originality’, ‘Significance’ and ‘Rigour’, and were instructed to apply their judgement in much the same way as they would for the REF. In addition, we asked them in this context to score each output on the same scale in terms of their transformative scope, though divided into two components: scope to transform social scientific enquiry, and scope to transform society (the latter being analogous to ‘impact’ potential).

The results indicate that on average, outputs from the TR scheme have a quality profile strikingly similar to those from Standard grants; slightly higher in fact, though the differences are marginal and not significant.

Figure 22: Review exercise – quality profiles of TR and Standard grants outputs (individual review level)



5 denotes a rating of ‘Very high’, 1 denotes a rating of ‘very low’. NB: excludes reviews where reviewers did not rate themselves as at least ‘fairly confident’

Likewise in terms of impact to have transformative effects on wider society, the TR outputs perform well, roughly on par with their Standard grant counterparts. This is significant because the TR scheme foremost pursues transformative research in the sense of transforming scientific enquiry itself rather than wider society. As there is no known relationship between these two dimensions of transformativity, it is worth highlighting that this is by no means an area of comparative weakness. Indeed, our interviews also highlighted that many TR grants have included outreach and stakeholder

engagement activities, and several have attracted interest from non-academic stakeholders. A small number have had highly significant levels of interest from policymakers and other organisations: 14 out of our 20 interviewees reported at least some level of policy or wider public interest.

It is furthermore important to note that the clearest difference in the quality profile between TR and Standard outputs is on the criterion of scope to transform social scientific enquiry or practice. The difference here is marginal, though it already suggests some level of qualitative difference that we consider in more depth in the following section.

Overall, there is clear evidence that the TR grants produce outputs that are entirely comparable in both quantity and quality to their counterparts in similar-sized ESRC standard grants. There are no grounds here to suppose that the TR grants are in any way unproductive or produce unsatisfactory outcomes on the terms of 'normal science'. We note of course that many outputs especially from the more recent TR grants (as well as from the more recently concluded ones in our comparator sample) are likely to still produce more outputs in the immediate future, so the comparative terms expressed here should be considered, rather than the absolute numbers, which are imperfect.

However, the critical question is whether these grants in fact go beyond satisfactory and high-quality outputs in the context of 'normal science,' and present any features that differentiate them from Standard grants to the effect that claims around heightened transformative character might be made. This is the subject matter of the following section.

6.1.1 A note on unsuccessful grants

An additional aim of Phase 2 of this evaluation was to assess whether there have been any projects that were unsuccessful and failed to meet their objectives. As our analysis of outputs listed on ResearchFish shows, 21% of TR grants listed no outputs at all – a comparable figure to the one yielded for our comparator sample of ESRC standard grants. However, among our interviewees of grant holders almost all, including of projects that have not listed outputs, noted that at least one or usually several outputs (mostly journal articles) were either pending publication, under review or in preparation. At the level of outputs and productivity, there is therefore at most a negligible number of grants that could be described as unsuccessful.

In qualitative terms, and on the notion of transformative research itself, there is likewise little evidence of unsuccessful projects: as we show below (see also Appendix J), large proportions of outputs were judged to be highly transformative on at least some criteria. Moreover, the vast majority of interviewees noted either that their research had either achieved transformative outcomes or was likely to do so in subsequent work. In two cases, grant holders did note that they did not see their project – or large parts of it – as being successful. This was attributed to the over-ambitious nature of their plan: risks had been taken in terms of embarking on uncertain research endeavours, and as the research progressed, the plans turned out not to be feasible. However, we stress that this is the case only for a very small number of projects, and that these need to be understood in the context of many others which had comparably ambitious scope and did in fact produce results.

6.2 The transformative character of the grants

Besides profiling and analysing the outputs from TR grants in more general terms, the key research question for this part of the evaluation was whether, to what extent and in what ways the outcomes and outputs from TR grants are in fact transformative, and whether they are in any way 'more' transformative than outputs in their Standard grant counterparts. Our findings in these respects are for the most part facilitated by our comparative outputs review exercise. Two observations are noteworthy here at the outset:

Firstly, our reviewers generally found it more challenging to rate outputs from the TR scheme: we allocated outputs to our six reviewers based on their stated areas of research expertise and interest, as well as our own internal judgement of the thematic and disciplinary substance of each output (the allocation was made by social scientists at Technopolis, PhD level or above). We then asked reviewers

to specify for each review how confident they felt about their judgement in terms of their expertise in each output's subject matter. 80 reviews were conducted respectively on outputs from the TR and Standard schemes. For 68 of the standard scheme outputs, reviewers rated themselves as 'fairly confident' or 'very confident' in their judgements, whilst for the TR outputs, this figure drops to 47. For many parts of our analysis, we subsequently found that the trends and relationships we highlight are more pronounced when reviews self-judged as 'less confident' or 'not at all confident' are excluded. These observations indicate, firstly, that judging the transformative scope of a piece of research requires a certain level of expertise, and secondly, that outputs from the TR scheme are likely somewhat harder to allocate and/or judge. Looking back to Phase 1 of this evaluation, the ESRC's use of a review panel, where a range of experts can come together to discuss merits of each application is once more merited by these findings.

Secondly, it was by far not only outputs from the TR scheme that were judged to be transformative overall – many outputs from the Standard grants received such verdicts as well. Besides scoring each output on a range of criteria pertaining to transformative scope as well as quality, reviewers were also asked to provide an overall judgement on whether they would describe each output as 'transformative'. The data resulting from this element suggest that transformative research is not limited to TR grants: the majority of outputs were in fact deemed overall transformative, although the proportion in the TR scheme is slightly higher.

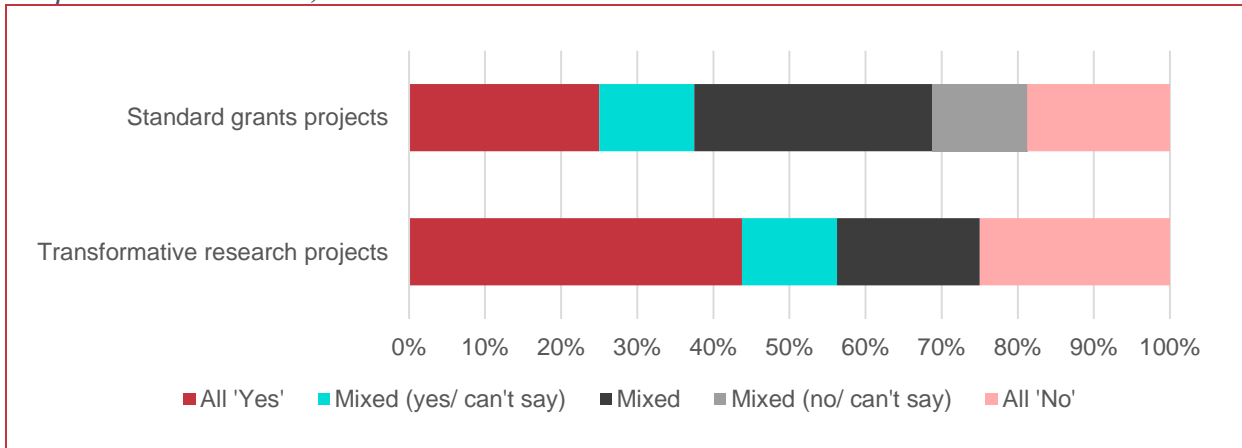
Table 16: Review exercise – verdict on transformative character (individual review level)

Overall, would you describe this output as 'transformative'?	TR		Standard	
	Count	%	Count	%
Yes	28	60%	37	54%
No	15	32%	24	35%
Can't say	4	9%	7	10%
Total reviews	47	100%	68	100%

NB: excludes reviews where reviewers did not rate themselves as at least 'fairly confident'

A picture more clearly in favour of the TR scheme emerges when we look not at individual reviews of outputs, but at entire grant projects. Where we had multiple reviews for the same grant (so either two reviews on the same output, and/or reviews on different outputs from one particular grant), we looked at the extent of consensus between reviews within projects. The resulting figures give a sense not of individual reviews, but of the extent to which there is consensus between reviewers around the project as a whole in terms of transformative scope. As Figure 23 shows, there are fewer Standard grant projects where our reviewers consistently found outputs to be transformative, though many more where some identified overall transformative scope whilst others did not.

Figure 23: Consensus and disagreement at the level of projects (question: Overall, would you describe this output as 'transformative?')



NB: based on 42 reviews covering 16 projects for the TR grants, 40 reviews covering 16 projects for the Standard grants – for each grant there were either 2, 3 or 4 reviews in total.

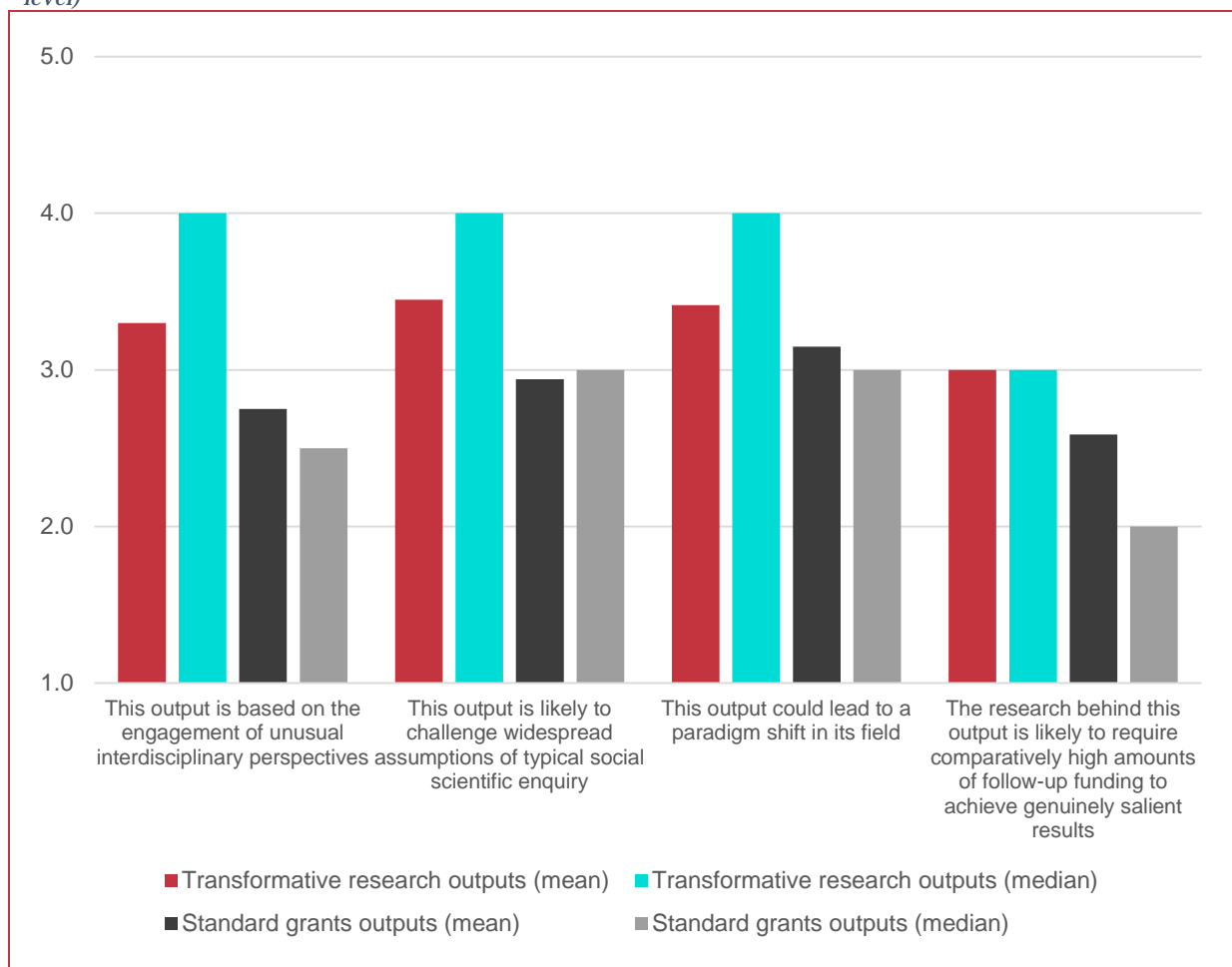
At the level of projects, but to a lesser extent also at the level of individual reviews of outputs, the TR grants are more commonly acknowledged as being overall of a transformative nature. However, many outputs of Standard grants are described as transformative too. In a sense, this says less about the TR scheme, and more about the perception that ESRC tends towards funding conservative projects – these data suggest this may in fact not be the case. It also reflects the problem that any funded research ought to make an original, 'new' contribution, which raises the question of where 'new' ends and 'transformative' begins.

6.2.1 Transformative qualities in detail

The overall question of whether an output or a project is transformative of course aggregates and potentially conflates a range of different possible criteria discussed in this report. On the individual criteria pertaining to possible definitions and characteristics of transformative research, the greatest differences between outputs from the TR scheme and Standard grants were on the following:

- “This output is based on the engagement of unusual interdisciplinary perspectives”
- “This output is likely to challenge widespread assumptions of typical social scientific enquiry”
- “This output could lead to a paradigm shift in its field”
- “The research behind this output is likely to require comparatively high amounts of follow-up funding to achieve genuinely salient results”

Figure 24: Review exercise – areas of greatest difference between TR and Standard grants (individual review level)



5 denotes ‘Agree strongly’, 1 denotes ‘Disagree strongly’. NB: excludes reviews where reviewers did not rate themselves as at least ‘fairly confident’

Smaller differences (usually in favour of the TR grants) also exist on several other criteria, but the differences are too small to interpret in a meaningful way. However, it is worth noting additionally, that we see more ‘peaks’ on several further criteria among the outputs of the TR scheme: The proportion of reviews that gave top ratings (so either ‘strongly agree’ or ‘very high’) is noticeably higher for the TR outputs in several categories.

Table 17: Peaks – percentages of reviews scoring the highest possible mark

	TR Scheme - % of reviews with the top mark (/5)	Standard grants - % of reviews with the top mark (/5)	Difference
This output presents a pioneering theoretical or methodological innovation	30%	13%	17%
This output represents a novel application of theory or methods to a new context	30%	18%	12%
This output is based on the engagement of unusual interdisciplinary perspectives	19%	6%	13%
This output is likely to challenge widespread assumptions of typical social scientific enquiry	17%	4%	13%

	TR Scheme - % of reviews with the top mark (/5)	Standard grants - % of reviews with the top mark (/5)	Difference
The research behind this output is likely to require comparatively high amounts of follow-up funding to achieve genuinely salient results	13%	3%	10%

NB: excludes reviews where reviewers did not rate themselves as at least ‘fairly confident’

Differences between TR grant and Standard grant outputs are tendencial rather than absolute. Yet, these data do suggest that TR grants tend to be more interdisciplinary, and more often appear to put established conventions and assumptions into question. It is especially noteworthy that the notion of paradigm shifts, drawn from Thomas Kuhn’s (1970) canonical work on transformative research (though in the context of physics rather than social science) is among the most evident areas of difference. The ambition to fund research that not only goes beyond the mainstream, but also challenges it, appears to be fulfilled at least to some extent in the TR scheme.

The heightened need for follow-up funding is likewise of interest: many of our interviewees confirmed that the grants enabled proof-of-concept research to take place, but rarely said that their ideas could be brought to their fullest possible fruition with a TR grant. These findings suggest that many TR grants are of ambitious scope, and that grant activities only mark the beginning of potentially extensive, long-term projects (if indeed they continue). On this particular criterion it should however be noted that the average figure is still squarely in ‘neutral’ territory, meaning that the evident need for further follow-up funding by no means applies to all grants, but it does apply more often or to greater extent than it does in Standard grants.

These findings are highly encouraging: although many outputs from Standard grants are also judged as being transformative, both overall as well as on the specific criteria measured in our review exercise, outputs from TR grants tend to have these qualities more often or to a greater extent, in particular in terms of unusual interdisciplinary qualities, and challenging or even paradigm-shifting potential. We also note once again that our interviews suggest many outputs are still in preparation or under review, so these could not be included in the exercise. There is a possibility that more ‘landmark’ publications may well be among them.

6.2.2 Transformative outcomes in context

Both our interviews with grant holders, as well as our outputs review exercise have highlighted that projects funded under the TR scheme have had many transformative outcomes. In the review exercise, this is evidenced in outputs scoring higher on average – as well as having more ‘peaks’ – on several possible characteristics of ‘transformative research’ than their counterparts in ESRC standard grants. In terms of impact, it is not possible at this point to contextualise this any further: the earliest publications from the scheme only date back to 2013 (most are far more recent) so a citation analysis or even a more qualitative assessment of how these outputs have been received or responded to by others would simply be premature at this point.

However, our interviews with grant holders give a more detailed sense of the ways in which some of these projects are transformative. As we discuss below in section 6.4, many of the projects are not seen by the grant holders as finite entities, but rather as the beginning of longer-term projects, a good number of which have already been extended beyond the TR grant itself. For this reason, substantive transformative effects are typically only emergent, for instance at a proof-of-concept level. In practice, this means that often a new research method or approach to a new field has been successfully trialled and results are being written up or recently published.

In some cases, however, even the short timelines so far enabled grant holders to point to genuinely transformative results from the grants. At this point these are generally less related to the impacts of published work, but to the process of conducting the research itself, or of early dissemination

activities. Critical for these early forms of transformative outcomes are most often unusual collaborations, which in themselves presented something new. Examples include:

- ‘Health of populations and ecosystems’: this project involved collaboration between health economics and environmental science, and at the same time brought together previously unacquainted groups of policymakers from the two domains. Greater integration and cross-coordination between these fields appears to be a result.
- ‘Beyond the banality of evil’: using simulation technology, this project succeeded in re-framing the Milgram experiments on obedience to authority figures (originally conducted by making participants believe they are administering electric shocks) in an ethically sound way. Dissemination and general public interest and awareness of the original experiments has raised considerable interest in this line of enquiry and the wealth of knowledge that may be gained from it.
- ‘Synchronous movement cooperation and the performing arts’: this project involved dancers in the design of Psychology experiments, the results of which are to be trialled in the NHS for the treatment of neuro-psychological disorders. This type of collaborative experiment design is in itself a novelty
- ‘Picturing the social’: this project led to the establishment of a research lab studying the sharing of images on social media around key world events. In itself, this presents a substantive outcome, but the profile of these activities was elevated when it received significant exposure after publishing a rapid response report to the viral picture of drowned Syrian boy Aylan Kurdi, using its newly developed capabilities.

We stress that these are the earliest possible signs of substantive transformations being achieved. Collaborations and exposure appear to be factors behind these most evident instances at this point. However, as we explain below, many other projects may well have significant impact in future, but the capacity for exposure is not such that instant results of the kind noted in the examples above are possible. Likewise, the projects noted above may also produce quite different transformative outcomes besides the early successes.

6.3 Determinants of transformative scope

6.3.1 *The importance of quality*

We further made use of our review exercise data to assess whether there are any noteworthy relationships between the individual criteria. Most notably, we find here that, reflecting considerations already made on the TR scheme’s assessment process, there is a strong link between quality and transformative scope. When we separate those review criteria that most directly resemble definitions of ‘transformative research’ (including the ESRC’s own) and our REF-style quality judgements, there are significant correlations almost across the board. Even in the ‘Rigour’ dimension, which intuitively appears less related to transformative scope than, say, ‘Originality’, there are significant (albeit weaker) significant correlations on all but one criterion. Likewise, our proxy for impact (‘scope to transform society’) correlates with all definitions of transformative research (in the scientific rather than social sense).

Table 18: Correlations between markers of quality and definitions of transformative research

	Originality		Significance		Rigour		Scope to transform society (Impact proxy)	
	<i>r-value</i>	<i>Strength</i>	<i>r-value</i>	<i>Strength</i>	<i>r-value</i>	<i>Strength</i>	<i>r-value</i>	<i>Strength</i>
This output presents a pioneering theoretical or methodological innovation	0.7117*	Strong	0.6145*	Strong	0.3769*	Moderate	0.4626*	Moderate
This output represents a novel application of theory or methods to a new context	0.6819*	Strong	0.5907*	Moderate	0.3470*	Moderate	0.4787*	Moderate
This output is based on the engagement of unusual interdisciplinary perspectives	0.4819*	Moderate	0.5091*	Moderate	0.2023*	Weak	0.4456*	Moderate
This research behind this output may generally be described as 'high-risk, high reward'	0.3717*	Moderate	0.3436*	Moderate	0.1849*	Weak	0.3048*	Moderate
This output is likely to challenge widespread assumptions of typical social scientific enquiry	0.5784*	Moderate	0.4575*	Moderate	0.152	Weak	0.3254*	Moderate
This output could lead to a paradigm shift in its field	0.6448*	Strong	0.6417*	Strong	0.2486*	Weak	0.4920*	Moderate
This output is likely to produce a broad base of knowledge, new thinking or insights	0.6951*	Strong	0.7774*	Strong	0.4385*	Moderate	0.6078*	Strong

All correlations in this table are positive (i.e. high scores on one criterion are associated with high scores on another). (*) denotes significance at 0.95. NB: excludes reviews where reviewers did not rate themselves as at least 'fairly confident'

These findings are especially significant as we instructed reviewers to use the quality-profiling in the same way as they would for REF-reviewing. As in our assessment of the TR scheme's application process, there is an evident connection between transformative scope on one hand, and conformity to established standards around quality (and indeed, wider impact) on the other. Fully exploring this relationship is a social scientific and theoretical task that goes beyond the scope of this evaluation. The most evident question would be whether transformative potential in itself signals quality and wider significance, or whether established markers of quality are preconditions for transformative scope of a piece of research to be pondered in the first place. In either case, these two dimensions are evidently connected.

6.3.2 The seniority-problem

This part of our analysis also highlighted a relationship between both quality and transformative scope on one hand, and issues around seniority on the other. We asked reviewers to note the extent to which outputs appeared to be typical of early career researchers as opposed to senior scholars. Reviewers' scores accurately identify more ECRs in outputs from the 2013-14 round than in the 2012-13 round (see section 4.1). At the same time, there are negative correlations between scores on this criterion and all markers of quality and definitions of transformative research.

Table 19: Negative correlations between early career research and quality/ TR scope

	'This output is more typical of early career researchers than of senior, established scholars'	
<i>Criteria</i>	<i>r-value</i>	<i>Strength</i>
This output presents a pioneering theoretical or methodological innovation	-0.4234*	Moderate
This output represents a novel application of theory or methods to a new context	-0.4473*	Moderate
This output is based on the engagement of unusual interdisciplinary perspectives	-0.4014*	Moderate
This research behind this output may generally be described as 'high-risk, high reward'	-0.3775*	Moderate
This output is likely to challenge widespread assumptions of typical social scientific enquiry	-0.3818*	Moderate
This output could lead to a paradigm shift in its field	-0.4597*	Moderate
This output is likely to produce a broad base of knowledge, new thinking or insights	-0.5421*	Moderate
ORIGINALITY	-0.5336*	Moderate
SIGNIFIANCE	-0.6262*	Strong
RIGOUR	-0.4685*	Moderate
TR SOPE - Social Science	-0.5242*	Moderate
TR SOPE – Society (Impact proxy)	-0.3727*	Moderate

All correlations in this table are negative (i.e. high scores on one criterion are associated with low scores on another). (*) denotes significance at 0.95. NB: excludes reviews where reviewers did not rate themselves as at least 'fairly confident'

In a sense, these figures can be expected, although the consistency of the relationships is nevertheless striking. The existence of specific ECR funding schemes both in the UK and internationally (e.g. ERC Starter Grants, Norwegian RCN Young Research Talents, Royal Society Research Grants) provides evidence that ECRs cannot readily compete with their more senior counterparts. The REF's concessions in terms of number of outputs to be submitted by ECRs highlight similar issues (see also Laudel & Gläser 2007). A tentative hypothesis in interpreting these data is that senior scholars will be more readily and immediately able to produce outputs that conform to established norms and quality than younger researchers with less track and experience. Given the links between these standards and transformative scope, it stands to reason that here too, ECRs are at a disadvantage. This is not to say that senior scholars are categorically more able to produce transformative research (our data allow for no such conclusion); however, our findings suggest that senior researchers may more readily be capable of translating their grants into impactful and significant outputs, whilst for ECRs these grants may be more of a starting point to begin work on novel subjects or approaches, where more significant outputs may only occur further down the line, as their scholarly ability and experience increases.

6.4 Further plans and hindsight-perspectives

In terms of satisfaction with the Scheme itself, our Phase 2 findings are consistent with Phase 1: Our interviewees almost categorically spoke very positively about their relationship with ESRC during the course of the grant period. The only noteworthy point in this respect was that seven out of our 20 interviewees noted that they had an extension granted. Several more noted they had requested one, but were denied. There is therefore a possibility that an extension of the timeframe of the grants (though not necessarily the amount, which was rarely problematised) could be contemplated. Notably,

many reported that the grant period tended to allow time for the research itself, but less often for preparation of publications, which often occurred after grants ended.

In terms of other difficulties around the grant, particularly pertaining to the transformative nature of the work, or other deviations from original project plans, we do not see any concerning results. Some grant holders noted minor changes to original plans, especially where unexpected opportunities arose (e.g. unexpected access to a particularly promising fieldwork site), but only two interviewees described changes as being significantly disruptive. There were also few reports of problems around publication: a small number of interviewees noted that publication of results was challenging, but this tended to relate to the interdisciplinary nature of their work, itself a well-established topic in the literature on the ‘science of science’ (Rafols et al 2012), rather than to its transformative character as such.

Beyond this, a further element we sought to assess in this phase of the evaluation were the potential and emerging long-term trajectories of the TR grants. Based in particular on our interviews, we find that in many cases the TR grants themselves mark only the beginning of longer-term endeavours, and that the notion of ‘proof-of-concept’ or early experimentation is a salient one. In terms of both transformative results, as well as the broader notion of success more generally, there was an overwhelmingly positive response to our interviews. However, many interviewees were at the same time cautious to note that successes or genuinely transformative results were emergent rather than complete. Notably, these results include grant holders whose grant ended more than a year ago, as well as the more recent completions of 2013-14 grants.

Table 20: Interviews – successful and transformative results

Interview question	Response summary (all numbers are out of 20)
Do you think your TR project was successful?	<ul style="list-style-type: none"> • 7 say enthusiastically yes, outright • 10 say yes, but that there is more work to do • 5 were a little more cautious • 1-2 were fairly despondent
Has your project been transformative?	<ul style="list-style-type: none"> • 16 say ‘yes’ in some form • 8 of these say yes, but it is ‘early days’ and in some ways too soon to tell • 4 tend towards ‘no’

Additionally, we find that the ‘publication trail’ from these grants tends to extend decisively beyond the end of grants themselves. Several interviewees noted in fact that whilst the grant period accommodated the research itself, it was not sufficient to include time to prepare publications. The sampling process for our review exercise confirmed this: though 2013-14 grants had largely also ended, comparatively fewer publications had been logged on ResearchFish for the second round. Likewise in our interviews, fewer grant holders from the 2013-14 call noted that they had had papers published yet. However, the majority of grant holders from the second round noted they had papers either under review or accepted, with most of the remainder noting that papers were either in progress or planned in the immediate future now that the grant had ended and time was available.

More strikingly, holders from the 2012-13 round also noted that they had publication in progress or under review, meaning that even over a year after grants ended, publication output was still occurring. In this context, it is worth noting that while our outputs review exercise as a whole showed greater transformative scope compared with Standard grants as demonstrated in the preceding sections, outputs from the 2013-14 call score consistently lower on average across our spectrum of criteria compared with those from the 2012-13 call (see our full review data in Appendix J). Based on the discussion so far, we attribute this to a number of factors:

- Projects from the 2013-14 call have only just ended, whilst those from the 2012-13 call ended over a year ago; the post-project publication trail is therefore less developed and fewer articles are already in the public domain available for review.
- Relatedly, earlier publications may be more experimental and less robust as they likely reflect the earliest work conducted within the grants. More broadly, our outputs selected for the 2013-14 call contain a larger share of working papers as opposed to journal articles.
- The higher incidence of ECRs in the 2013-14 call may also be at play here: these grant holders may have been less able to readily produce high-quality outputs at the same rate as their counterparts.

In terms of longer-term future trajectories, it is likewise clear that funded projects do not end with the end-point of the grant. Having largely acknowledged that there is 'more to be done', most interviewees had plans to continue their transformative research in some form. Although the lack of designated follow-up funding to the TR scheme was often lamented in our data collection for both Phase 1 and Phase 2 of this evaluation, there are in fact many substantive efforts and results in this area. Out of our 20 interviewees:

- 6 have secured further funding from a range of sources (incl. ERC, ESRC and other funders)
- 8 have grants applications in progress or under review
- Others have either been unsuccessful and seeking to apply elsewhere, or do not have clear plans yet.

Notably, in terms of successful as well as on-going efforts to secure further funding, there appears to be little difference between grant holders from the two calls. These findings further strengthen the sense that ESRC has not simply funded fixed 18-month projects with a finite and predictable range of outputs, but that new research agendas have frequently been set in motion, some of which had large degrees of success right away in terms of both academic output as well as wider interest, but many of which will likely continue to grow as more publications emerge, and further funding is secured to take the initial findings to fuller fruition. These longer-term trajectories, of which TR grants were often noted as the critical first step, are especially noteworthy when we consider the extent to which grant holders deem the TR scheme as essential to kick-starting their transformative activities. We asked our interviewees what would have happened if they had not been able to secure their TR grant. Out of 20 interviewees:

- 5 said outright that their project would not have happened as nobody else would have funded it
- 5 said their project was unlikely to have happened without TR call for the same reasons
- 5 said it would have gone ahead but slower or with reduced scope, relying on smaller funding pots or reducing research proposals to the less 'transformative' elements of the project
- 3 said they would have tried to obtain the same level of funding elsewhere
- 2 were unsure/ unclear

On a final note, it is worth pointing out that this sense of long-term trajectories mitigates to a large extent the finding that a higher proportion of ECRs in the later calls of the scheme may have meant a lower incidence of outputs that immediately show signs of outstanding quality and transformative scope. As Phase 1 showed, the scheme evolved not least to become less biased towards senior, established scholars. If the ESRC's intention was to fund discreet projects with immediate or short-term high quality results but little further development, then our Phase 2 findings might suggest that the shift away from senior academics has been problematic. However, given that there are clear signs of longer-term research agendas – that would moreover have been impossible or improbably without the TR scheme – the scheme's evolution towards a more diverse range of grant holders in terms of seniority ought to be welcomed.

7 Conclusion and recommendations

7.1 Taking the Scheme forward

Our evaluation of the ESRC's Transformative Research Scheme shows that it has been a success in terms of identifying and funding research projects with transformative scope, and producing long-term research agendas and individual outputs on a suitable scale that are often more transformative than Standard grant counterparts on several dimensions. The Scheme manages to attract a pool of applications that generally contain a high level of novel methodological and/or theoretical approaches, whilst also carrying key hallmarks of contemporary professional academia, i.e. suitable for publication in established journals, ethically sound and related to topics of broad interest. The combination of anonymous review and the Pitch-to-peers workshop successfully draws from this pool those that are judged by our applications review exercise to have the strongest range of transformative characteristics.

The two assessment stages perform distinct tasks. The sift panel seeks out and debates transformative scope, whilst also creating a level playing field through the anonymity component. The Pitch-to-peers events have a greater focus on ensuring suitability of the applicant, de-risking, checking scholarly quality, as well as using the opportunity to ask for clarifications, which projects of a transformative nature are especially likely to require.

However, this is not the full story. The Scheme has involved a long-term learning exercise, in which deliberation, dialogue and a range of different feedback loops have led to an increasingly functional process, where institutions gain increasing understanding about which candidates to put forward, assessment panellists reflect continuously on their own creeping conservatism and how best to negotiate and identify the boundaries of what 'transformative' may mean, and where panel chairs and organisers learn how to manage and optimise the process.

Within the Scheme, a growing culture of transformative research is evident; this is to some extent feeding back to institutions, most evidently through successful applicants as well as in some cases the institutional payments from the 2012-13 call. The extent to which this culture of transformative research has trickled through to the wider social science community cannot be judged from this evaluation.

An abrupt end to the Scheme is not desirable. However, there may be a need for the Scheme to be better embedded in the wider social science funding landscape; not only to enable this learning exercise to gain wider traction, but also because the stand-alone nature of the Scheme currently may pose problems for the research projects themselves. Some grant holders have secured further funding, and many more have applications under review or plan to prepare applications in the future. However, the projects funded have a tendency to be ambitious and present only the beginning of longer-term agendas (which is very much in the spirit of the Scheme), and whilst a small-scale grant is a preferable option in terms of managing overall risk, the lack of immediately obvious channels for follow-on funding may be a point of concern.

7.2 Funding Transformative Research: A re-appraisal

'Transformative research' in contemporary social science is a complex and multi-faceted notion. Our analysis suggests that it can involve pioneering theoretical or methodological innovations or engagement of unusual disciplinary or interdisciplinary perspectives. It may fall into the high risk/high reward category, present a genuine challenge to established conventions, and may be of large scope necessitating large amounts of follow-up funding to come to full fruition. However, no single one of these characteristics is a definitive singular marker. The notion of paradigm-shifting ideas that challenge established 'mainstreams' and ways of doing things is likewise prominent (though not essential). But whilst this suggests some continuing relevance of Kuhn's (1970) analysis of research, the notion of unsuitability to present-day professionalised academia is not borne out: transformative and potentially paradigm-shifting ideas are judged as such if they are suited to academic publication,

contain scope for wider impact and, are ethically sound and more broadly conform to established markers of research quality. That said, it should be noted that whilst our findings do not preclude the possible emergence of research so revolutionary that systemic resistance from academic peers and research management becomes a long-term obstacle, this does not appear to be a salient phenomenon around this Scheme.

Given the pluralistic meaning of transformative research, suitably identifying and funding it necessitates above all the long-term provision of a space for incremental learning, discussion and understanding, where the meaning of the term, ways of identifying it, its boundaries and its challenges, can be suitably established. Feedback loops between grant holders, applicants, research funders, institutions and assessment panellists are critical here to ensure both a suitable learning curve within that space, as well as dissemination around the social science community are made possible. Identifying individuals able to take the lead in such learning processes is critical, especially with regard to senior assessment panellists (e.g. panel chairs), who are able to counter the creeping conservatism of assessment procedures, as well as the potential for bullish competition and unnecessarily tactical scrutiny of ideas that are by their very nature embryonic and thus potentially still vulnerable.

Transformative research not only takes many forms – it can also come from many different places. We find no significant patterns in terms of seniority (i.e. early career researchers or high-level professors), gender or prestige of host institution. Particularly as early career researchers may lack the experience of their senior counterparts to immediately produce high-quality outputs, a pluralistic space for researchers to apply and potentially become involved is critical – the Scheme’s anonymous review process is a helpful facilitator in this sense.

However, pressure on funders and the need to control the inevitably heightened level of risk associated with transformative research also requires scrutiny. Moreover, transformative ideas easily incur a greater need for questions, clarifications expositions than more conservative research proposals. For both these reasons, and indeed to cement the notion of a space for learning about transformative research more generally, the Pitch-to-peers events present an apt approach.

7.3 List of recommendations

Based on various elements of our findings, we conclude here with a range of recommendations pertaining to various aspects of the Scheme, from its operational optimisation to its wider importance and salience.

7.3.1 *Headline recommendation: the Scheme should continue*

- The ESRC has successfully created a space in which a growing culture and understanding of transformative research is taking hold. Though modifications are possible (and in some cases necessary, as described below), this space needs to be preserved in order for this process to continue. The involvement of individuals with an interest in transformative research and/or past experience of the scheme is likewise important here.

7.3.2 *Operation*

- We find little evidence to suggest that the amount of funding per grant is problematic. However, there is scope to recommend a slight extension of the timeframe (e.g. from 18 to 24 months).
- It is worth clarifying and codifying the rules of engagement for the Pitch-to-peers workshops: introductions; use of technology; expected minimum number of questions; help with timekeeping for presenters are currently conducted in a far from uniform manner, which may make presentation easier in some groups than in others; how to address conflicts of interest, especially at the Pitch-to-peers events, but also at the sift panel meetings;
- The ten-point marking scale was subject to some debate and is clearly being used differently by different assessors and fellow applicants. As this is the standard scale across many different ESRC funding tools, it is essential for comparative purposes and cannot be replaced as such. However,

clearer guidelines for how to use it in this particular Scheme would likely be beneficial. This might for instance involve introducing a number of different scales (e.g. one for transformative scope, one for scholarly quality, one for risk), which then combine to give a score out of ten;

- Some of the £50k institutional payments have clearly been used for activities that contribute to a growing culture of transformative research at institutions (e.g. 'sandpit' events, symposia, mini-grants, teaching and training activities). However, lack of specificity or oversight of how these payments are used casts doubt on whether there has been overall value for money. If such institutional payments are re-introduced, this should involve a clear plan of what to use it for, and more stringent accountability and transparency;
- The potential for discrimination at Pitch-to-peers events needs to be addressed. Although panel chairs and organisers have developed the events over time and alleviated many of the concerns highlighted around the first event in 2013, issues such as presentation anxiety, or overnight travel causing obstacles for instance with regard to disability or childcare need to be taken into account.

7.3.3 *Integration*

- Universities have generally managed to ensure that applications well suited to the Scheme are submitted. However, they do so based on fragments of intelligence often gained from former applicants. A narrower and hence clearer definition of transformative research is hardly feasible. However, case studies or profiles of funded projects, as well as the findings of this evaluation, should be circulated to universities to help the selection process;
- The self-reflective and engaged role of panellists is critical to the proper functioning of this Scheme. Panellists need therefore to be especially carefully selected. A system of recommendation or application specifically for this Scheme is an option worth considering. Recruitment of former grant holders to the panel has proven especially helpful, so this is worthwhile continuing and expanding.

7.3.4 *Beyond the Scheme*

- It is worth considering identification or introduction of clearer pathways to follow-up funding channels (or fast-tracking mechanisms to existing big-grant schemes), to be clarified to applicants from the outset. Eligibility could be made dependent on a minimum profile of outputs and outcomes (e.g. 'proof of concept');
- The anonymous review format has wide appeal. Extension to other schemes should be considered. However, though this approach addresses some important challenges (notably bias of peer review and path dependency of academic careers), it does not constitute an adequate tool for de-risking and cannot function as the sole component of a funding tool. Should this element of the Scheme therefore be applied to other areas of funding, second component aimed at de-risking and checking applicants' scholarly ability is likely to be necessary;
- The two-page application format might likewise be useful elsewhere, especially in schemes that could benefit from a significantly accelerated application process.

Appendix A Details of methodology

A.1 Survey

We conducted an online survey of all applicants to the 2012-13 and 2013-14 calls of the Scheme. Details of responses are noted below.

Table 21: Survey – Respondent Details

Respondent Details (n=81; N=132; response rate: 61%)		
<i>Gender</i>		
	Response Per cent	Response Count
Male	60.5%	49
Female	38.3%	31
Other	1.2%	1
<i>Age at the time of application</i>		
	Response Per cent	Response Count
under 30	0.0%	0
30-39	28.4%	23
40-49	34.6%	28
50-59	25.9%	21
60+	11.1%	9
<i>Seniority at the time of application</i>		
	Response Per cent	Response Count
Research fellow (incl post-doc)	4.9%	4
Lecturer	14.8%	12
Senior lecturer	9.9%	8
Reader	9.9%	8
Professor	59.3%	48
Other academic position	1.2%	1
Not based at an academic research institution	0.0%	0

A.2 Interviews

To follow up on the initial findings of the survey, and to produce deeper qualitative information about the procedural elements of the Scheme, as well as its challenges and future possibilities, we conducted a series of 20 telephone interviews, each lasting around 30 minutes. We aimed to interview members of each of the following groups, in addition to a small selection of other external experts:

- A selection of grant beneficiaries: to help generate a clearer and more detailed picture of the Scheme from the academics' own point of view, and as such presenting the most direct continuation and deepening of the survey results;

- A selection of research officers for social and economic science at beneficiary universities charged with selection of funding applications to put forward: this is a critical group that plays a decisive role in determining, which applications are forwarded to research funding councils;
- Assessment panel members and chairs: these individuals need to be interviewed in order to provide the ‘opposing’ perspective, as well as a wider view on how the Scheme worked, whether there were operational challenges and/ or opportunities, and how the notion of ‘transformative research’ materialised in the commissioning process;
- ESRC Committee members and ESRC Officers: this group can provide input on the wider strategic benefits and challenges of the Scheme.

Table 22: List of interviewees

Type	Name	Position	Institution	Interview date	Interviewer
ESRC	Jeremy Neathey	Deputy Director of research	ESRC	05-06-2015	Peter Kolarz
ESRC	Julie McLaren	Head of Society and Global Security Group, Sponsor of 2013/14 TR call	ESRC	05-06-2015	Peter Kolarz
Grant holder	Alison Liebling	Professor of Criminology and Criminal Justice	University of Cambridge	08-06-2015	Kalle Nielsen
Grant holder	Dawn Watkins	Senior Lecturer, School of Law	University of Leicester	26-05-2015	Kalle Nielsen
Grant Holder	Jonathan Halket	Lecturer, Department of Economics	University of Essex	27-05-2015	Kalle Nielsen
Grant Holder	Nathaniel Tkacz	Assistant Professor, Warwick Centre for Interdisciplinary Methodologies	University of Warwick	02-06-2015	Kalle Nielsen
Grant holder / Panellist	Eric Herring	Reader in International Politics	University of Bristol	09-06-2015	Peter Kolarz
Grant Holder / Panellist	Hilary Graham	Professor of Health Sciences	University of York	10-06-2015	Peter Kolarz
Panellist	Claire O'Malley	Professor of Learning Science	University of Nottingham	27-05-2015	Martin Wain
Panellist	Simon Collinson	Dean of Birmingham Business School	University of Birmingham	27-05-2015	Martin Wain
Panellist	Richard Thorpe	Pro Dean for Research	Leeds University	11-06-2015	Kalle Nielsen
RO officer	Andrew Clark	Director of Research Planning	University College London	03-06-2015	Martin Wain
RO officer	Patrick Lansley	Research Development Manager	University of Bristol	29-05-2015	Martin Wain
RO officer	Rachel Baker	Research Development Officer	University of Cardiff	29-05-2015	Martin Wain
RO officer	Sharron Pleydell-Pearce	Divisional research Facilitator	University of Oxford	26-05-2015	Martin Wain
RO officer	Sue Coleman	Senior Research Support Advisor	University of Edinburgh	28-05-2015	Martin Wain
RO officer	Michael Woods	Head of the New Political Geographies Research Group	University of Aberystwyth	13-07-2015	Martin Wain

Other	Paula Bailey	Senior Leaders integrators Manager	EPSRC	12-05-2015	Peter Kolarz
Other	Michael Stampfer	Managing Director	WWTF, Austria	21-04-2015	Peter Kolarz
Other	Terttu Luukkonen	Chief Advisor	ETLA, Finland	16-04-2015 (e-mails only)	Peter Kolarz

The exact interview script was adapted to fully suit each of the groups, notably for those interviewees who began as grant holders but had acted as panellists subsequently. However, questions centred largely on the following points for all:

- What is the general understanding of respondents of the notion of ‘transformative research’? Why does it necessitate a distinct funding scheme?
- The extent to which transformative ideas can be pitched to funders with normal peer review processes (as opposed to schemes ring-fenced for transformative work)
- How challenging is internal competition in the universities to be allowed to be one of the applicants to this Scheme, and how is the selection made?
- What happens during each part of the commissioning process, and does the process, or particular parts of the process, aid/hinder the funding of transformative research?
- Have any design/operational/procedural challenges arisen, and (how) has the Scheme been amended to overcome these challenges?
- What are the implications of the ‘two applications only from Top 11 Institutions’ and ‘two applications from all Research Organisations’ models?
- [For 2012-13 grants only] How were the additional funding sums of £50k per institution used? Did this encourage transformative research, and what lessons can be learned
- What are the key strengths and weaknesses of the scheme?
- Should the ESRC continue to run the transformative research scheme as a separate call? If so how can it be strengthened?
- Should the ESRC replicate aspects of the scheme in other funding streams to encourage more transformative research elsewhere? If so, how could this be done?

A.3 Observation

As an additional step towards assessing the design, operational and procedural aspects of the Scheme, we included an observational element. As the third round of funding on this scheme was underway during the evaluation, we were able to attend critical events occurring as part of this process, specifically:

- The sift panel short-listing meeting on 22-04-2015 in London;
- The ‘Pitch to Peers’ event and the Commissioning Panel meeting on the 17 and 18 June 2015 in Birmingham.

A.4 Project review

For this exercise, a panel of seven peer reviewers was asked to review a total of 93 2-page proposals submitted to the scheme over the three calls. All proposals of funded projects, as well as random samples of 24 shortlisted for Pitch-to-peers workshops and 24 rejected at review stage (stratified equally across the three calls) were each reviewed by two peers (allocation based on expertise). The

proposals were fully anonymised and reviewers were not informed about the actual outcome in the Scheme of each proposal.

Table 23: Sample of proposals for the review exercise

	2012/13 Call (N=67)	2013/14 Call (N=69)	2014/15 Call (N=106)	Total for review (N=242)
Rejected (Random sample)	8	8	8	24
Shortlisted (Random sample)	8	8	8	24
Accepted (All)	20	13	12	45
Total	36	29	28	93

For each proposal, both reviewers had to fill in a template (attached for reference in Appendix A.5.2 , involving:

- 17 statements and a 5-point likert scale (ranging from ‘disagree strongly’ to ‘agree strongly’, as well as a ‘cannot judge’ option);
- A final verdict, asking whether they consider the proposal to be transformative (yes / no / cannot say);
- Instructions to explain their verdict;
- Space for additional comments.

In twelve out of the 93 proposals, at least one of the two reviewers had noted subsequently that they did not feel able to score the proposal in question, or checked the ‘cannot judge’ option on several statements. For these twelve proposals an alternative reviewer was identified, in order to obtain two reviews for all proposals based on the largest possible degree of expertise.

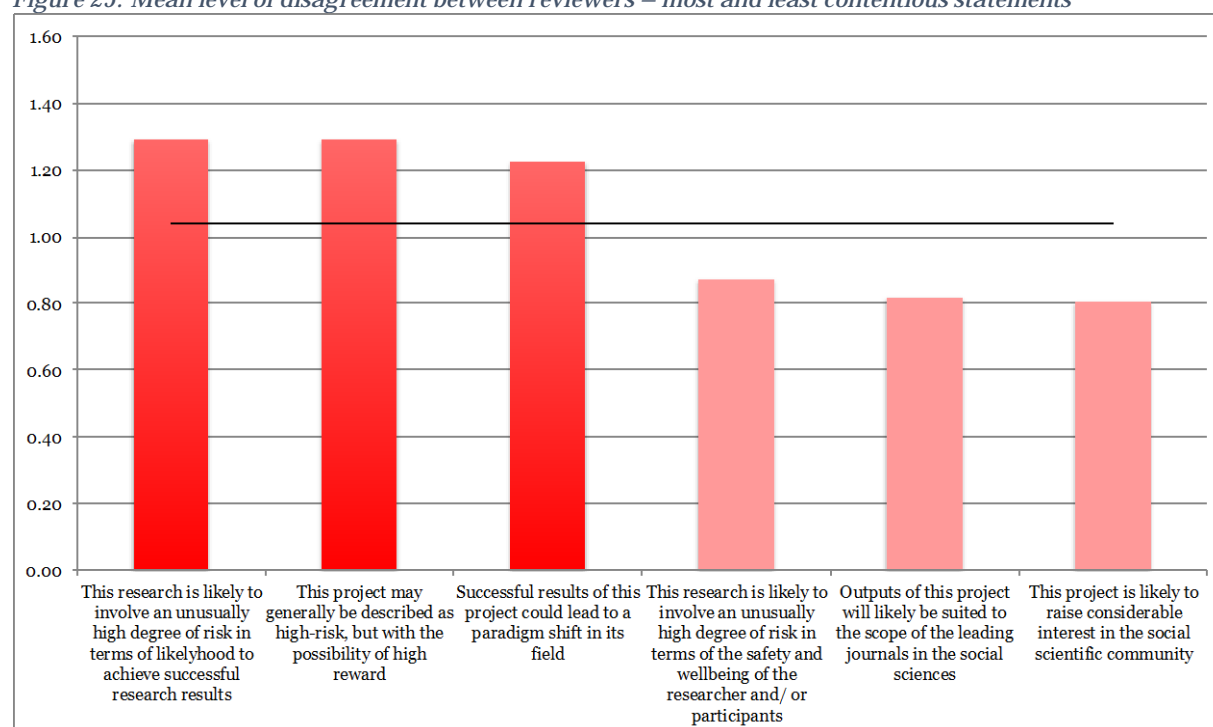
A.5 Divergence

In peer reviews seeking to lead to decisions on funding or publication, it is critical for reviewers to reach agreement on scores. However, in the absence of this, such agreement was not necessary here. In fact, observation at the sift panel and pitch-to-peer events strongly suggested that disagreement can be a key hallmark of transformative proposals. As such, the divergences between reviewers for each proposal were an important point of investigation in itself.

Overall, the level of disagreement between reviewers was low. Given 93 proposals and 17 questions, the exercise yielded 1581 pairs of judgements on a 5-point likert scale. Reviewers only disagreed by more than two points in 8.1% of possible cases, and by more than one point in just 27.3% of cases. On average, reviewers were therefore almost exactly one point (1.039) apart in their rankings, with complete disagreements very rare.

Out of the 17 main template statements, those relating to risk in terms of the feasibility of the research were the most contentious, i.e. on which reviewers tended to disagree the most, as was the question of whether a project presented a potential paradigm shift. Those statements on which reviewers generally had the strongest consensus were around research ethics, as well as suitability to leading academic journals and overall interest that a project might raise in its field. However, even these most and least contentious areas of the data do not deviate to an unexpected extent from the overall levels of disagreement seen between reviewers across the exercise.

Figure 25: Mean level of disagreement between reviewers – most and least contentious statements



*Black line denotes the overall mean level of disagreement of 1.039

Final scores for each proposal were reached by taking the mean between the two reviewers on each criterion. Following re-reviewing of proposals where a reviewer had noted unfamiliarity with the topic, there were a total of 27 instances (1.7% of possible cases) where the ‘Cannot judge’ option had been selected, excluding the final template question (‘Overall, would you describe this proposal as transformative?’). In those cases, no mean was taken, and the single score given by one reviewer was used for further analysis.

A.5.1 Reviewers

Table 24: List of reviewers for the project review exercise

Name	Institution	Discipline (sub-discipline)	Expertise Overview	Also participated for the Phase 2 review of outputs
Prof Rupert Brown	University of Sussex	Psychology	Social psychology; Intergroup relations and group processes. Discrimination, Prejudice and prejudice reduction; Acculturation processes; Hate crime; Social identity processes.	No
Prof John Carpenter	University of Bristol	Policy: Social work and applied social science	Social Work and Social Policy sub-panel for REF2014 - anything within that domain policy and services disability, mental health, children and families, learning disability, psychotherapy pedagogical research	Yes

Prof Susan Condor	Loughborough University	Psychology: Social and political psychology	Social psychology Political psychology History of Science Mixed methods/ qualitative methods Prejudice, stereotyping, racism and discrimination Self and identity Nationalism and national identity Public attitudes and public opinion; Ambivalence and ideological dilemmas; Social psychological aspects of citizenship; Political inaction and apathy.	Yes
Prof Graham Crow	University of Edinburgh	Sociology: Methods, interdisciplinary work, theory, households, families	Sociology, Social Policy (but au fait with all social science disciplines) Topics: Rural, Urban, Community, Family, Social Change Methods: qualitative, mixed, comparative, ethics	Yes
Prof Andrew Jones	City University London	Geography: Economic geography/ interdisciplinary	Human geographer Economic and urban globalization of the economy and transnational firms as global organizations. Finance and business services. Global city networks, creative industries work migration / mobility. Sociology of work voluntary work global workers in overseas development.	Yes
Prof Martin Laffin	Queen Mary University of London	Business/ management: Public management, public policy	Public sector management and public policy includes service delivery (involving private and voluntary sectors), UK devolution, regionalism and local government. Housing policy managing under austerity	Yes
Prof Wendy Larner	University of Bristol	Geography: Human geography and sociology (Moved to Victoria University of Wellington, NZ by the time the Phase 2 review happened)	Human geography globalisation neoliberalism governance regulation co-production social policy economic development strategies gender activism	Yes

A.5.2 Template

ESRC Transformative Research

Project Appraisal Exercise

Project number (3 digit code on the application's file name):

Project Title:

Reviewer name:

Please rate the project on each of the following criteria:

	Disagree Strongly	Disagree somewhat	Neutral	Agree Somewhat	Agree Strongly	Cannot judge
This project presents a pioneering theoretical or methodological innovation						
This project represents a novel application of theory or methods to a new context						
This research is based on the engagement of unusual disciplinary and interdisciplinary perspectives						
This research is likely to involve an unusually high degree of risk in terms of likelihood to achieve successful research results						
This research is likely to involve an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants						
This project may generally be described as high-risk, but with the possibility of high reward						
This project is likely to challenge widespread assumptions of typical social scientific enquiry						
Successful results of this project could lead to a paradigm shift in its field						
This project is likely to raise considerable interest in the social scientific community						
Outputs of this project will likely be suited to the scope of the leading journals in the social sciences						
This project may struggle to find favour with peer reviewers in its aligned disciplines						
This project lends itself to eventual application of results (if successful) outside of academia						
This project is likely to require comparatively long lead times to achieve successful results						
This project is more typical of early career researchers than of senior, established scholars						
Even if successful, results of this project may struggle to gain widespread acceptance in the academic community						
If successful, this project is likely to produce a broad base of knowledge, new thinking or insights						
This project is likely to require comparatively high amounts of follow-up funding to achieve successful results						

Overall, would you describe this project as ‘transformative’?

Yes	
No	
Can't say	

Please explain your answer (briefly):

--

Please add any further thoughts, qualifications or comments you feel might be important:

Appendix B Bibliography

- Arnold E, Luukonen T, Boekholt P, Nooijen A, Jävolrka Z and Zuijdam F (2013) *Evaluation of the Academy of Finland*. Helsinki: Ministry of Education and Culture.
- Chubin DE and Hackett EJ (1990) *Peerless Science: Peer review and US science policy*. Albany: SUNY Press.
- ESRC (2013) 'Top 11 Institutions by Research Expenditure', in: ESRC 2012/13 Annual Report; available: http://www.esrc.ac.uk/images/ESRC_12_13_AR_tcm8-27176.pdf
- Grant J and Allen L (1999) 'Evaluating High Risk Research: an Assessment of the Wellcome Trust's Sir Henry Wellcome Commemorative Awards for Innovative Research.' *Research Evaluation*, 8(3): 201-4.
- Häyrynen M (2007) *Breakthrough Research: Funding for High-Risk Research at the Academy of Finland*, 8/07, Helsinki: Academy of Finland.
- Horrobin DF (1996) "Peer review of grant applications: a harbinger for mediocrity in clinical research?" *The Lancet*, 348: 1293-5.
- Kuhn TS (1970) *The Structure of Scientific Revolutions*, (2nd edn), Chicago: Chicago UP.
- Lakatos I and Musgrave A (1970) *Criticism and the Growth of Knowledge*, Cambridge: CUP.
- Lariviere V, Ni C, Gingras Y, Cronin B and Sugimoto CR (2013) 'Bibliometrics: Global gender disparities in science.' *Nature*, 504(7479): 211-13.
- Laudel G and Gläser J (2014) 'Beyond breakthrough research: Epistemic properties of research and their consequences for research funding.' *Research Policy*, 43(7): 1204-16.
- Laudel G and Gläser J (2007) 'From apprentice to colleague: The metamorphosis of Early Career researchers'. *Higher Education*, 55(3): 387-406
- Luukkonen T, Stampfer M and Strassnig M (2015) 'Evaluation practices in the selection of groundbreaking research proposals.' *WWTF Working Paper*: 13-08-2015. Available: http://wwtf.at/other_activities/erc_expert_group/index.php?lang=EN
- Luukkonen T (2012) 'Conservatism and risk-taking in peer review: Emerging ERC practices.' *Research Evaluation*, 21(1): 48–60.
- Mertens DM (2008) *Transformative Research and Evaluation*, Guildford: Guildford Press.
- National Science Board (2007) *Enhancing Support of Transformative Research at the National Science Foundation*. Washington DC, National Science Board.
- Oakley A (1998) 'Gender, Methodology and People's Ways of Knowing: Some Problems with Feminism and the Paradigm Debate in Social Science.' *Sociology*, 32(4): 707-31.
- Rafols I, Leydesdorff L, O'Hare A, Nightingale P and Stirling, A (2012) 'How journal rankings can suppress interdisciplinary research: A comparison between Innovation Studies and Business & Management.' *Research Policy*, 41(7): 1262-1282.
- Roy R (1985) 'Funding Science: The Real Defects of Peer Review and an Alternative to It', *Science, Technology, and Human Values*, 10(3): 73–81.
- Travis GDL and Collins HM (1991) 'New light on old Boys: Cognitive and Institutional Particularism in the peer review system.' *Science, Technology & Human Values*, 16(3): 322-41.

Wagner CS and Alexander J (2013) 'Evaluating transformative research: A case study of the NSF Small Grants for Exploratory Research programme.' *Research Evaluation*, 22(3): 187-97.

Wessely S (1998) 'Peer review of grant applications: what do we know?' *The Lancet*, 352: 301-5.

Appendix C Statistical analysis of Phase 1 applications review data

This note explains the statistical analysis conducted to:

1. Analyse what ‘transformative research’ might mean based on the relationship between the 17 different definitions provided to reviewers from Technopolis’ project review exercise (herein, the reviewers) to assess whether a project was ‘transformative’ or not (sub-section 1.1) and the reviewers’ final verdict (sub-section 1.1);
2. Estimate the extent to which the PRE reviewers were able to predict the assessment (in terms of scoring) made by the (i) Pitch-to-peers panel, (ii) Pitch-to-peers peer and (iii) sift panels (sub-section 1.2);
3. Estimate the extent to which the PRE reviewers were able to predict the final decision on projects (in terms of projects being accepted, shortlisted or rejected) (sub-section 1.2).

We find that there is a high level of agreement among reviewers, who provide the same verdict for a project in 79.6 per cent of the cases.

Additionally, we find that there are four definitions that are positively and strongly related to the final verdict. If we look at project level, Definition 3 (engagement of unusual disciplinary and interdisciplinary perspectives) and 17 (high amounts of follow-up funding needed to achieve successful results) seem to predict very well the probability of arriving to positive (un-coordinated) consensus among reviewers, while Definition 1 (pioneering theoretical or methodological innovation) and 6 (high-risk, but with the possibility of high reward) seem to be the most important factors to explain the probability to vote yes, at individual level. This suggested that those aspects are the main features that define what ‘transformative research’ means in the eyes of reviewers.

Furthermore, We find that reviewers are good at predicting the scores of the Pitch-to-peers panel when they arrive to a positive (un-coordinated) consensus (i.e. both give a final positive verdict a the project). We did not find a similar relationship when looking at the scores from the Pitch-to-peers peer and sift panels, or when looking at the data from the reviewers at individual level (instead of at project level).

Finally, there is some evidence to suggest that reviewers’ scores are able to predict final results. We find that the composite index is positive and statistically significant to the indicator that shows that a project has been accepted or shortlisted. More specifically, we find that a high aggregate scoring (composite index generated using factor analysis and including the 17 definitions) is correlated with a higher the probability of the projects being accepted or shortlisted. The effect disappears when we exclude the 11 projects that were accepted even though they did to pass the 7 score threshold at the sift panel stage. This latter result means that reviewers are actually better predicting the final actual result.

The following sub-sections provide more details on the methodological approach and findings.

C.1 What does ‘transformative research’ mean?

In order to assess the various possible dimensions of what ‘transformative research’ might mean the study team asked reviewers to assess a list of 93 projects that had been submitted to the scheme, and that had been accepted, shortlisted or rejected. The project review exercise included all the accepted projects plus a random selection of shortlisted and rejected projects.

Each reviewer had to provide scores from 1 to 5 (where 1 means strongly disagree and 5 means strongly agree) for 17 questions each of which contained a definition of ‘transformative research’.

Two reviewers were assigned per project. Our analysis, focus mainly on the score per project, unless stated otherwise. The score per project is obtained by estimating an average of the scores provided by each reviewer per project.

We first investigated whether or not there was common understanding of what ‘transformative research’ mean across reviewers. We test this by looking at the degree to which reviewers tend to agree in on their final assessment of the projects. They are asked: ‘overall, would you describe this project as ‘transformative’? Reviewers have the option to reply: ‘yes’, ‘no’, or ‘cannot say’.

We find that there is a moderate level of agreement among reviewers. They are in agreement with their verdict in 42 per cent of the cases, as shown (Table 25).

Table 25: Verdict

	Verdict	Frequency	Percentage
1	No/No	13	14.0
2	No/Cannot say	5	5.4
3	Yes/No	35	37.6
4	Yes/Cannot say	14	15.1
5	Yes/Yes	26	28.0
		93	100

Source: Project review exercise

Moving into the analysis of the 17 definitions, we estimate the extent to which certain definitions are highly related to the final verdict provided by the reviewers. Is a high a score in some definitions highly correlated with a positive final verdict?

We estimate a regression using the variable verdict as a dependent variable (i.e. variable that we want to explain) and a selection of the 17 questions as our explanatory variables. Note that we use a selection of the definitions instead of the full set as adding a long list of indicators that are somehow correlated as explanatory variables could decrease the explanatory power of the model.

Following standard process, we first eliminate outliers by removing the scores that are 2 standard deviations above or below the mean, for each of the 17 definitions.

We then select the relevant set of definitions based on an initial partial correlation analysis: we analyse which definitions are strongly correlated with final verdict. This initial analysis identifies 5 definitions, which will used as explanatory variables:

- Definition 3. This research is based on the engagement of unusual disciplinary and interdisciplinary perspectives;
- Definition 5. This research is likely to involve an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants;
- Definition 9. This project is likely to raise considerable interest in the social scientific community;
- Definition 12. This project lends itself to eventual application of results (if successful) outside of academia;
- Definition 17. This project is likely to require comparatively high amounts of follow-up funding to achieve successful results.

Given that our dependent variable (final verdict) is categorical (i.e. not continuous) we cannot estimate a linear regression, which requires having a continuous indicator as a dependent variable. Instead we need to rely on a probabilistic model. We opted for creating a binary variable that takes the value of 1 if the final verdict has been ‘Yes/Yes’ or ‘Yes/Cannot say’, i.e. if at least one of the reviewers thought that overall the project is ‘transformative’.

After estimating the probabilistic model, **we find that there are four definitions (out of the five) that are statistically significantly related to the final verdict.** Definition 3, 9 and 17 are positively correlated to the final verdict. This means that a positive verdict tends to come accompanied with a high score in those three questions.

On the other hand, Definition 5 is negatively correlated to the final verdict. This means that a positive verdict tends to come accompanied with a low score in this question. Definition 12 is also negatively correlated but this result is not statistically significant.

Based on the estimated model we predict the probability of a positive verdict based on two potential scores: 1 and 5 (Note that this is a friendlier way to present the results coming from the econometric regression)¹². The results are shown in (Table 26) **They mean to say that the (predicted) probability of a positive verdict is:**

- 88.9 per cent if Question 3 had a score of 5, and only 0.6 per cent if the question had a score of 1;
- 88.7 per cent if Question 9 had a score of 5, and only 0 per cent if the question had a score of 1;
- 94.2 per cent if Question 17 had a score of 5, and only 3.9 per cent if the question had a score of 1;
- Only 1.1 per cent if Question 5 had a score of 5, and only 61.5 per cent if the question had a score of 1.

Note that in the table only some of the results are statistically significant (shown by the stars given to their level of statistical representativeness). Taken in conjunction with the prior finding, this means that the three questions are indeed statistically significantly correlated to the final verdict (as we already explained) but that the individual predictions at each score level (prepared to help the interpretation of the results) sometimes are not, and this is mainly due to the existence of few data points at these score levels.

Table 26: Predicted probabilities of a positive verdict, given value of scores

Scores	Predicted probability	Std. Err.	P> z	Level of statistical representativeness 1/
Definition 3				
1	0.01	0.013	0.637	
5	0.89	0.080	0.000	***
Definition 9				
1	0.00	0.000	0.859	
5	0.89	0.083	0.000	***
Definition 17				
1	0.04	0.063	0.536	
5	0.94	0.087	0.000	***
Definition 5				
1	0.62	0.100	0.000	***
5	0.01	0.029	0.711	

1/ Level of statistical representativeness: * p<0.05, ** p<0.01, *** p<0.001. Based on 89 observations at project level (excludes outliers)

¹² We estimate this for each of the four definitions individually, keeping all other definitions constant at their mean values.

Similar results are obtained for a more ‘restrictive’ definition of a positive verdict. When estimating the regression to compare those projects in which both reviewers gave a final verdict of yes, versus the rest of the projects, we find that Definition 3 and Definition 17 remain statistically significantly correlated to the final verdict.

We also ran a similar exercise at reviewer level rather than at project level. That gives us 186 observations rather than 93.

In this case a different set of definitions become statistically significant. This is probably because there are different definitions that are relevant to predict (un-coordinated) consensus (yes/yes) and a different set of questions that are relevant to predict individual overall verdicts. **At individual level we found that there are two definitions are statistically significantly related to the final verdict.** Two questions are positively correlated, these are:

- Definition 1: This project presents a pioneering theoretical or methodological innovation;
- Definition 6: This project may generally be described as high-risk, but with the possibility of high reward.

Note that in these estimations we also controlled for the ‘reviewers’ effect, i.e. we add variables that take the value of 1 if a project have been assess for a particular reviewer to control by the fact that each of them have a different tendency to provide positive verdicts (for instance, one reviewer gave a positive final verdict 70 per cent of the times while other did it only 33.3 per cent of the times).

Similarly to what we did in the prior exercise we use our estimated model to predict the probability of a positive verdict based on two potential scores: 1 and 5. **They mean to say that the probability of a positive verdict is:**

- 89.0 per cent if Question 1 had a score of 5, and only 0% if the question had a score of 1;
- 75.9 per cent if Question 6 had a score of 5, and 19.3% if the question had a score of 1.

Table 27: Predicted probabilities of a positive verdict, given value of scores

Scores	Predicted probability	Std. Err.	P> z	Statistical representativeness 1/
Definition 1				
1	0.000	0.000	0.754	
5	0.890	0.044	0.000	***
Definition 6				
1	0.193	0.088	0.029	***
5	0.759	0.083	0.000	***

1/ Level of statistical representativeness: * p<0.05, ** p<0.01, *** p<0.001. Based on 168 observations at individual level (excludes outliers)

C.2 Understanding of ‘transformative research’ across panels and reviewers

One interesting issue to analyse is the extent to which the reviewers from the project exercise review are able to predict the results from other panels.

We find that reviewers are not able to predict results individually, however, they are good at predicting the scores of the Pitch-to-peers panel when they arrive to a positive (un-coordinated) consensus (yes/yes). Results for the econometric exercise are shown in Table 28.

The estimates are based on a linear regression (given that the scores provided by each panel are continuous). We estimate three different regressions, one per panel, and use the indicator verdict as an explanatory variable for all of them (this is a binary variable that takes the value of 1 if there is positive (un-coordinated) consensus). We also add indicators to capture the ‘call’ effect, as the average panel scores have increased over time/call. An indicator for “2012-13” is excluded as one year has to remain as a base year (i.e. a fixed point against which the other years are compared to in the regression)¹³. The figure shows that only the coefficient for the indicator verdict is statistically significant in the first regression, the one that uses the score Pitch-to-peers panel as a dependent variable.

Table 28: Regression analysis (Ordinary Least Squares), *at project level*

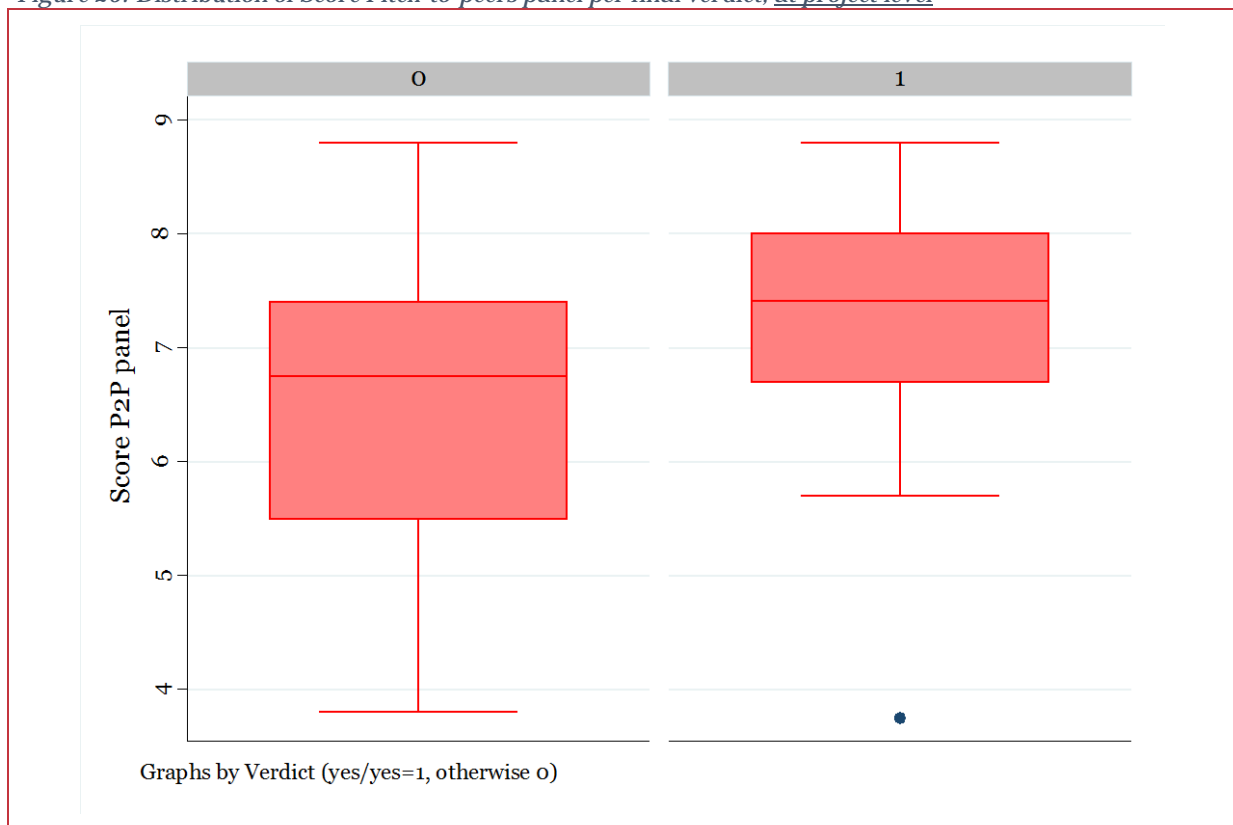
	Score Pitch-to-peers panel (coeff/se)	Score Pitch-to-peers peer (coeff/se)	Score sift panel (coeff/se)
Explanatory variables			
Verdict Yes/Yes (=1)	0.852*	-0.0220	-0.178
	(0.360)	(0.308)	(0.326)
2013-14 call (=1)	0.671	0.903*	-0.130
	(0.404)	(0.345)	(0.368)
2014-15 call (=1)	0.350	1.119**	0.632
	(0.402)	(0.344)	(0.374)
Constant	6.176***	5.978***	6.676***
	(0.293)	(0.251)	(0.265)
Number of observations	60	60	75
R-squared	0.114	0.186	0.057

Standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001

Figure 26 presents a simple graph that reinforces the finding for the regression analysis. It shows a graph with the distribution of the Pitch-to-peers panel scores across two sets of projects, those with a final positive verdict (group 1) and the rest (group 0). The graphs shows the data divided in quartiles. The boxes show the area cover for the data between the lower (25th) and the upper (75th) quartiles, while the line that cross the boxes represent the median. The lines extending vertically from the boxes indicate variability outside the upper and lower quartiles and the dot represent an outlier. It shows that the projects with a positive verdict are clustered around high scores provided by the Pitch-to-peers panel.

¹³ The regression excludes the negative Verdicts (no/no and no/cs) as by definition we should expect a different direction in the results (a negative association), hence including those verdicts would render our results meaningless.

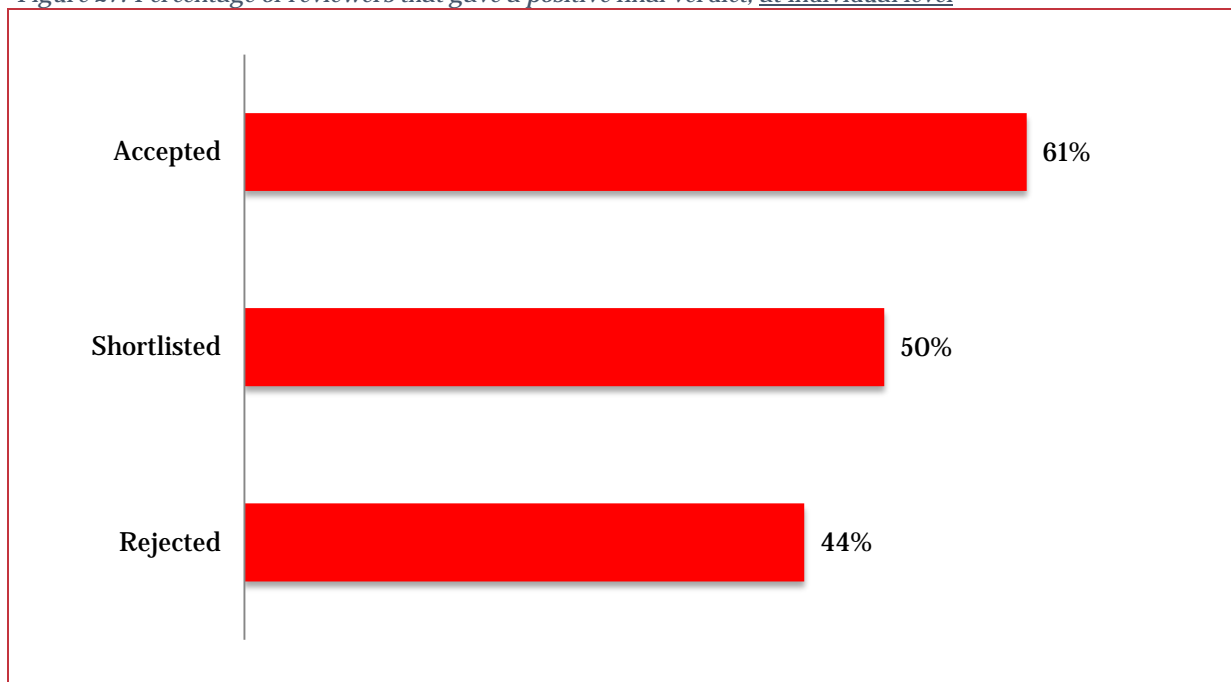
Figure 26: Distribution of Score Pitch-to-peers panel per final verdict, at project level



Finally, we analysed whether or not the reviewers were able to predict the final result and we find that **there is some evidence to suggest that reviewers' scores are able to predict final results.**

Unfortunately, we cannot run a regression to compare the reviewers' verdict with the final decision, as they are both categorical variables. Probabilistic regressions cannot cope with two binary variables, one as dependent and the other as explanatory variable. We can see from a simple graph (Figure 27) that the projects that were accepted were the ones for which the majority of reviewers provided a positive verdict (61 per cent). However, a relatively high percentage of reviewers also provided a positive verdict for projects that were rejected or shortlisted.

Figure 27: Percentage of reviewers that gave a positive final verdict, *at individual level*



Based on 186 observations

To further analyse these relationship through an econometric estimation, we created a composite index that groups the 17 definitions into a single (continuous) indicator, using factor analysis¹⁴.

Finally, we estimate a probabilistic regression that uses the final funding decision as dependent variable and the composite indicator as explanatory variable. We use three definitions for the final funding decision. Each definition/indicator receives the value of 1 if:

- The project has been accepted, and zero otherwise;
- The project has been accepted or shortlisted, and zero otherwise;
- The project has been shortlisted, zero if it was rejected (this excludes accepted projects).

We find that the composite index is positive and statistically significant for the first two indicators (i.e. accepted versus shortlisted or rejected and accepted or shortlisted versus rejected). In other words, we find that a high aggregate scoring (composite index) is correlated with a higher probability of a project being accepted or shortlisted. **In order words, reviewers do tend to provide overall high scores to those projects that will be accepted or shortlisted, which provided some evidence that they are good at predicting the final result.** The results are shown in Table 29.

The effect disappears when we exclude the 11 projects that were accepted even though they did to pass the 7 score threshold at the sift panel stage. This latter result means that reviewers are actually better at predicting the final actual result.

¹⁴ Using a probabilistic estimation we find that this single indicator is positively correlated to a positive verdict (and this relationship is statistically significant).

Table 29: Regression analysis (Probit), at individual level

	Indicator 1: The project has been accepted, and zero otherwise	Indicator 2: The project has been accepted or shortlisted, and zero otherwise	Indicator 3: The project has been accepted, zero if it was shortlisted (this excludes <u>rejected</u> projects)	Indicator 4: The project has been shortlisted, zero if it was rejected (this excludes <u>accepted</u> projects)
Reviewers' score (Composite indicator)	0.381*** (0.114)	0.260* (0.117)	0.344* (0.135)	0.0498 (0.150)
2013-14 call (=1)	-0.190 (0.231)	-0.128 (0.246)	-0.125 (0.281)	-0.00262 (0.318)
2014-15 call (=1)	-0.267 (0.242)	-0.137 (0.259)	-0.298 (0.287)	0.0328 (0.328)
Constant	0.0824 (0.157)	0.724*** (0.170)	0.505** (0.189)	-0.0261 (0.229)
Number of observations	173	173	127	90

Standard errors in parentheses;

* p<0.05, ** p<0.01, *** p<0.001

C.3 Additional notes on methods and findings

For the task of cross-referencing Scheme outcomes with the review data, a less-than-clear-cut picture initially emerged. A direct replication of Scheme outcomes and judgement of our review exercise was never anticipated: whilst our exercise purely sought to understand and judge transformative scope, the assessment process of the Scheme itself pursues other concerns in parallel, notably scholarly quality, fit with the ESRC's remit and adequacy of impact statements. Effectively, the review data was used to assess whether transformative scope was nevertheless a demonstrable predictor of outcome in the Scheme, despite the necessity for assessment panellists to also consider other factors, and if so, *what kind* of transformative scope.

Assessing correlations between each of the 17 characteristics and all possible scores and outcomes at the various stages of the Scheme, no individual strong correlations were found in any particular instance. However, when we consider more pluralistic understandings of 'transformative research', a combination of factors acts as a very strong predictor of outcome.

Appendix D Comparator schemes

Table 30: Details of other transformative research funding schemes

Organisation (Country)	Name of scheme	Basic details	Characteristics/ evaluation
EPSRC (UK)	Bright IDEAS Award: the Big Pitch	<ul style="list-style-type: none"> Grants of up to £250k Duration: 18 months 	<ul style="list-style-type: none"> Limited to a single investigator Peer review is performed anonymously Next stage: 'dragon's den' style pitch event, 15 mins presentation, 30 mins questions 3 Physical sciences and 8 engineering calls to date High degree of risk expected All researchers qualified to apply for EPSRC funding are eligible Exists alongside several other measures for transformative research, chiefly networking and discussion events: IDEAS Factory; Sandpits; Creativity@home
ERC (EU)	Advanced grants	<ul style="list-style-type: none"> Grants of up to € 2.5m per grant (in some cases up to € 3.5m per grant) Duration: up to 5 years 	<ul style="list-style-type: none"> Proposals should involve new, ground-breaking or unconventional methodologies, whose risky outlook is justified by the possibility of a major breakthrough with an impact beyond a specific research domain or discipline Targets researchers who have already established themselves as independent research leaders in their own right Sole evaluation criterion: scientific excellence of researcher and research proposal (transformative capacity is assumed to be implied to some extent if the right individuals are given suitable resources)
ERC (EU)	Starting grants	<ul style="list-style-type: none"> Grants of up to € 1.5m per grant (in some cases up to € 2m per grant) Duration: up to 5 years 	<ul style="list-style-type: none"> Eligibility: 2-7 years of experience since completion of PhD Aimed at addressing insufficient opportunities for young researchers to develop independent careers. Recognises link between young researchers and new ideas/ potential for paradigm shifts Proposals are evaluated by selected international peer reviewers: excellence is the sole criterion. Both the research project and the Principal Investigator are evaluated
NIH (USA)	Director's pioneer award programme	<ul style="list-style-type: none"> Awards are for US\$500k Direct Costs each year for five years 	<ul style="list-style-type: none"> To be considered pioneering, the proposed research must reflect ideas substantially different from those already being pursued in the investigator's laboratory or elsewhere. Investigators at all career levels are eligible Proposals evaluated through NIH peer review system, but reviewers emphasise: <ul style="list-style-type: none"> Significance and innovation of proposed project, The investigator (ability and track of innovation), Suitability of proposal to the award mechanism (sufficient risk/ potential impact/ significant new direction) Summary of evaluation findings available¹⁵

¹⁵ <https://dpcpsi.nih.gov/sites/default/files/CoC-051413-Pioneer-Award-Program-DP1.pdf>

Organisation (Country)	Name of scheme	Basic details	Characteristics/ evaluation
NSF (USA)	NSF Small Grants for Exploratory Research (SGER)	<ul style="list-style-type: none"> Ran from 1990 to 2006 Replaced by EAGER and RAPID Grants of up to US\$200k (average of 54k) over 2 years 	<ul style="list-style-type: none"> Emerged from concerns around peer review blocking transformative research Minimal set of guidelines Proposals often emerged through discussion between programme manager and researcher Approval at discretion of programme manager; full bypass of peer review Significantly less than overall allocated funds were distributed in grants, but success rate was high. Consultation with programme managers is a key reason for this. Evaluation available (Wagner & Alexander 2013)
Research Council of Norway	Young Research Talents	n/a	<ul style="list-style-type: none"> Objective of this scheme includes 'promote scientific renewal and development of disciplines and/or generate new knowledge about issues relevant to society' Researchers with between 2 and 8 years of experience post-PhD are eligible Grants awarded through peer review Criteria include 'Boldness in thinking and scientific renewal', denoted by: <ul style="list-style-type: none"> Bold hypotheses High potential for significant theoretical advancement Original methodology Creative approach to expanding the current knowledge base in the field
Wellcome Trust (UK)	Sir Henry Wellcome Commemorative Awards for Innovative Research	<ul style="list-style-type: none"> 1-year awards of up to \$50k Later increased to 18 month awards of up to £85k 	<ul style="list-style-type: none"> Initial outline proposal of 2 pages; a selection of these (judged most innovative) were then invited to submit a full application. Early career researchers encouraged to apply, but all are eligible Evaluation finds this scheme indeed funded research that was deemed more innovative, risky and adventurous than regular grants (Grant & Allen 1999).
NSERC (Canada)	Discovery Frontiers	<ul style="list-style-type: none"> CAS1m per year over 4 years 	<ul style="list-style-type: none"> Discovery Frontiers grants support a limited number of large international activities, opportunities or projects that are of high priority in the context of advanced research in Canada. No annual competition cycle. A limited number of proposals for large undertakings with a high level of impact and visibility will be accepted in each competition. Proposals should be for concerted, broad-based, collaborative efforts and not individual research projects Initiatives need to be 'high risk/ high pay-off; transformative/ disruptive; paradigm-shifting or accelerate establishment of expertise in an emerging field 2-stage application process <ul style="list-style-type: none"> Letter of Intent and 5-page proposal + 2-page team overview reviewed by international review committee to decide on suitability to the scheme Full application then requested and peer reviewed on a broader range of criteria

Appendix E Observation notes: Sift Panel 22-04-15

The meeting was attended by 4 representatives from ESRC, as well as the assessment panel, which was composed of 18 members (including one chair and three sub-chairs), though a small number of panel members were absent. They had notified the chair and arrangements had been made for other panel members to read their review notes and present them. Each attendee had been given a list of all reviewed proposals, sorted by mean score (highest to lowest), with conflicts of interest noted next to projects where applicable.

The Chair began by briefly introducing the purpose of the meeting and facilitating introduction of all attendees, and checking whether all conflict of interest forms had been duly completed and submitted.

A 5-minute presentation then outlined the purpose of the scheme, the definition of transformative research (ESRC's definitions + acceptance of risk element), as well as a brief explanation of the current evaluation and the purpose of Technopolis' presence at the event. The point about accepting/rewarding risk whilst ensuring proposals were nevertheless of legitimate and feasible scope was particularly emphasised.

Past panellists were invited to make any potentially helpful comments before proceedings began. Two members noted that panellists should not be too harsh, and to remember that this was the first stage of two: if there were doubts about essentially promising proposals, they recommended to try and shortlist, so that the Pitch-to-peers event could be used to 'hear more'.

Out of 104 proposals, the aim was stated to select around 30 (absolute maximum 36) for the Pitch-to-peers events, which would then aim to fund about 15 (absolute maximum 20). The chair noted that there may well be difficult decisions to make on the margins (a 10-point scale was used throughout, 7 denoted the cut-off point).

The approach was that on each proposal, the first assessor would present their view, then the second assessor was invited to add if they wished, and the panel was to agree on the verdict, with the possibility for other panellists to add further points. Over the course of the day it became apparent that most panellists had read at least a small portion of proposals other than the ones they themselves had been asked to review.

The list of proposals had been divided into four chunks:

- Those that had received good scores from both reviewers;
- Those that had received diverging scores with an overall good mean score;
- Those that had received diverging scores with a lower mean score;
- Those that had received low scores from both reviewers.

Proposals in the latter category were not to be discussed, but rejected outright. As such, there were a total of 47 proposals to be discussed, though the top-3 from the latter category were ultimately brought back in for discussion, to check whether these were genuinely beyond all scope for shortlisting.

A panel member noted at this point that mean scores were a problematic point of reference: transformative research was likely to divide opinion in some form. The chair acknowledged this and noted that panel members should be aware of this possibility over the course of the day.

Finally, a panel member enquired how exactly conflicts of interest had been determined without compromising anonymity. The chair explained that conflicts of interest had been identified based on institutional attachment, but that this could relate to the PI or members of partnering institutions. This appeared to instil confidence that anonymity had been ensured as much as possible even when assessing conflicts of interest.

The issue of how to identify and reward transformative research was raised many times, both implicitly and explicitly:

- One proposal deemed not especially transformative but of exceptionally high quality was approved;
- The point about transformative content being more important than quality and robustness was made again by the chair after the first few had been reviewed;
- Panellists noted that it was difficult to strike a balance between risk and credibility;
- Another panellist noted that this meeting would not determine definitive funding outcomes, so it was 'ok to be a bit more lenient';
- "As soon as we unpack something, we begin to talk it down" (Panellist);
- "A key criteria should be, 'do we want to hear more?'" (Panellist);
- The problematic distinction between 'new/novel' and 'transformative' was noted on a few occasions: all funded research ought to add something new in some form – so where is the cut-off line? There was no clear answer to this question;
- After the lunch break, the chair reaffirmed that the issue of divergence between assessors' scores should be taken seriously, and urged panellists to consider as much as possible the merits of proposals that had received one good and one bad score;
- Some time later, the chair expressed frustration that none of the proposals in the third group (diverging scores, overall sub-pass mark) had been shortlisted;
- Overall, the chair did an excellent job at clarifying scope and purpose of the meeting, ensuring due process (conflicts of interest, ensuring all were satisfied with agreed scores);
- The panellists likewise had clearly made efforts to consider their proposals, as well as some that they had not been asked to review. There was due deference to panel members with superior knowledge of particular topics of proposals;
- In general, younger panel members seemed more critical and were usually quicker to point out methodological or other quality-related problems and shortcomings, whilst more senior members appeared more easily able to reflect on the merit of ideas;
- It was fairly clear that the 10-point scale was not used in the same way by all panellists, with some openly admitting they only marked as high as 7 as a rule, and others noting they gave the highest possible marks to encourage transformative ideas. Yet, much was made of the scale (to shortlist, scores had to be at least 7, and discussions of borderline-proposals needed to be followed by a decision on whether or not to increase the mean score to 7);
- 25 proposals were ultimately graded 7 or above and therefore shortlisted for the Pitch-to-peers event. This included 6 that had been given a 'question mark' status during the discussion as there was no conclusive decision. As there were plenty of places left, all were then subsequently included for shortlisting.

Table 31: Sift Panel – Thematic summary of comments and discussion points

Totals: All proposals (50)	Transformative scope	Proposal quality/scholarly aspects	Impact/wider significance	Risk	Ethics	TOTALS
Total comments (incl. from non-assessors)	67 (73)*	50 (53)	30 (32)	14 (16)	7 (8)	168 (182)
Total supportive comments (incl. from non-assessors)	44 (45) 66% (62%)	15 (15) 30% (28%)	24 (26) 80% (81%)	3 (4) 21% (25%)	0 (0) 0% (0%)	86 (90)
Total critical comments (incl. from non-assessors)	23 (28) 34% (38%)	35 (38) 70% (72%)	6 (6) 20% (19%)	11 (12) 79% (75%)	7 (8) 100% (100%)	82 (92)
PAIRS: Total disagreements (incl. from non-assessors)	13 (15)	6 (6)	3 (4)	1 (2)	0 (0)	23 (27)
Percentage of proposals which triggered disagreement	26% (30%)	12%	6% (8%)	2% (4%)	0%	

Table 32: Sift panel – Comments on successful proposals

Totals: Successful proposals (26)	Transformative scope	Proposal quality/scholarly aspects	Impact/wider significance	Risk	Ethics	TOTALS
Total comments (incl. from non-assessors)	40 (45)*	23 (25)	21 (23)	11 (12)	5 (6)	100 (111)
Total supportive comments (incl. from non-assessors)	29 (30)	8 (8)	17 (19)	3 (3)	(0)	57 (60)
Total critical comments (incl. from non-assessors)	11 (15)	15 (17)	4 (4)	8 (9)	5 (6)	43 (51)

Table 33: Sift panel – Comments on rejected proposals

Totals: Rejected proposals (24)	Transformative scope	Proposal quality/scholarly aspects	Impact/wider significance	Risk	Ethics	TOTALS
Total comments (incl. from non-assessors)	29 (30)*	27 (28)	9 (9)	3 (4)	2 (2)	70 (73)
Total supportive comments (incl. from non-assessors)	15 (15)	7 (7)	7 (7)	0 (1)	(0)	29 (30)
Total critical comments (incl. from non-assessors)	14 (15)	20 (21)	2 (2)	3 (3)	2 (2)	41 (43)

* Where other panellists made additional comments (aside from brief endorsement of what assessors said), these were additionally noted. Figures in brackets denote number of comments including those made by people other than the assessors who had marked the proposal in question.

Appendix F Observation notes: Pitch-to-peers 17/18-07-15

F.1 Panel briefing

The event began with a panel briefing. This included first an overview of the purpose and scope of the meeting by ESRC staff and the Chair. Whilst the briefing itself contained little of note [having a high-level ESRC person present to explain the purpose of the scheme would have been helpful all-round], subsequent discussion yielded a few points:

- Over the course of the meeting, about half the room made points, contributions or asked questions;
- Gender balance and age range seemed fairly adequate;
- It was agreed that after each presentation, panellists would ask questions first in order to cover all issues raised in the initial panel review phase. Only after that should fellow applicants ask questions;
- The criteria for the presentation were re-stated as: Fit with social science remit, transformative content, creativity, ability of applicant to carry out the research, impact of the research. It was noted that applicants are fully aware of the criteria;
- The chair pointed out that applicants should not be judged based on their performance as a presenter;
- The chair also noted it was important to make applicants understand that applicants are not competing against each other, but against the criteria. Chair and a former PI both noted the importance of this;
- The question was raised of what fellow applicant's scores should be for. It was noted that these should be used in cases where panel scores disagree, to help come to a decision;
- The issue of deciding by mean rather than median scores was noted [again], though to little effect;
- The issue of conflicts of interest was raised: what if a panellist turns out to know an applicant [impossible to tell prior to this point]. This was not dealt with especially well – it is indeed unclear how this might be dealt with at all in this context;
- The chair raised the issue of ensuring a comfortable atmosphere more generally: it was deemed critical to make the event as non-competitive and non-stressful as possible. One panellist compared it to a PhD viva.

F.2 Presentations

- Prior to the presentations, the chair addressed the panellists and applicants. She noted that the scheme came about due to the perception of the ESRC as being risk-averse and focusing excessively on applicants' track;
- A few people looked quite nervous;
- Unsure whether the set-up of cabaret-tables had much merit;
- Session 1, Group 1: 6 males and 9 females in the room; session 2, group 3: 9 males, 6 females; session 3, group 2: 7 females, 8 males;
- Group 3 had introductions of everyone at the start. They were the only one to do this;
- One presenter in group 1 had a laptop to read from. She apologised for this, but it was deemed acceptable (so unclear what the exact rules are);
- In group 1, presenters were introduced by the sub-char, in group 2 they introduced themselves;
- Some had a stand-alone presentation; a few framed their presentation in direct response to the questions asked in first round feedback.

F.3 After the presentations

F.3.1 De-brief

Following the presentations, the Chair held a de-briefing with all applicants and panellists, this yielded a few points:

- Assessment criteria (the 10-point scale) were noted as being too fine-grained. 10 points were too much to reliably classify proposals;
- Shorter scales, but for several different criteria were mooted as an alternative, so e.g. a 3-5 point scale for TR / quality / other aspects;
- Adding a confidence level to judgements was also noted as a possibility, as attendees felt they could judge some presentations much better than others;
- The chair noted that comments had generally focused on risk [reflecting our own indications that the purpose of Pitch-to-peers is in fact de-risking];
- In terms of innovative/ transformative character, it was felt that there was a strong emphasis on methodological rather than theoretical novelty;
- 7 mins presentation and 15 mins for questions was generally deemed tough, but very helpful and a good way of clarifying the project. With slightly more presentation time, it was felt that issues raised in question-time could have been pre-empted, but reserving 15 mins for questions was seen as essential in either case, and should not be shortened;
- The venue was judged to be excellent. Apparently much better, more relaxed and nicer than last year's in Swindon;
- The social event on the previous evening was deemed a useful occasion, though it was noted that it should not be called 'networking', or anything that implies a distinct task – applicants were unsure as to how to behave towards each other;
- The complete ban on technology in presentations was met with some criticism. In practice this was indeed variously well enforced, so while absence of PowerPoint was welcomed, some use of laptops or phones might be an option for the future;
- Overall, this de-briefing and the Chair's extensive questions and requests for verbal feedback and discussion makes it a very self-reflexive scheme.

F.3.2 Panel meeting

- At the start of the panel meeting, the issue of conflicts of interest was raised again. It was decided that if a panellist had a conflict of interest on a particular application, they should simply be quiet and not contribute to the discussion;
- Applicants who scored highly (comfortably above the pass mark), and those that were decisively below the pass mark were not discussed (or only very briefly, i.e. 'everyone happy to fund this? – yes.' Where marks were divided between pass and fail, or where panel and applicant scores diverged, there was discussion. Notable examples:
 1. The proposal was judged to be a great idea, but that it needed a bigger grant and was not feasible on this small grant scheme. Panellists asked whether it would be possible to 'forward' this application to a scheme of larger grants, but it was not;
 2. The applicant was deemed to have suffered a disadvantage due to a language barrier, which weakened his presentation and subsequent discussion. Their proposed project also had some feasibility problems, and it was overall unclear whether they had a proper understanding of the subject-matter;
 3. One panellist had given an extremely low score, while all others were favourable. The panellist went to great lengths to explain that this application was of very low quality and convinced

everyone else eventually. It appeared to be a case of one panellist having genuine expertise and all others making judgements on limited knowledge of the field;

4. The application was considered very strong in terms of the idea, though poor on the project and approach itself;
 - a. It was noted in this context that the project was transformative in terms of society, but NOT in terms of social science itself – which was deemed the more significant dimension!
 5. The application was a very mixed bag of high and low scores. All elements of the application were discussed and there was no consensus. The Chair and others proposed that since nobody could rule out that it might be very fruitful, it should be 'given a shot';
- Finally, the chair thanked panellists, and praised the excellent level of collegiality.

Appendix G Comparing Panellists' and fellow applicants' Pitch-to-peers scores

Table 34: Comparing Panel and peer scores for the 2012/13 call

	Panel mean	Peer mean	Panel rank	Peer rank	Funded	Funded (peers)*
Project 1	8.75	8.20	1	1	Y	Y
Project 2	8.50	7.50	2	3	Y	Y
Project 3	8.50	7.50	3	4	Y	Y
Project 4	8.50	3.90	4**	29	Y	N
Project 5	8.33	6.80	5	9	Y	N
Project 6	7.67	6.10	6	12	Y	N
Project 7	7.50	5.90	7	14	Y	N
Project 8	7.33	7.44	8	5	Y	Y
Project 9	7.33	5.20	9	24	Y	N
Project 10	7.25	5.60	10	18	Y	N
Project 11	7.00	7.67	11	2	Y	Y
Project 12	7.00	5.78	12	16	Y	N
Project 13	7.00	5.11	13	26	Y	N
Project 14	7.00	4.80	14	27	Y	N
Project 15	6.75	5.90	15	15	Y	N
Project 16	6.75	5.30	16	21	Y	N
Project 17	6.50	5.30	17	22	N	N
Project 18	6.50	3.11	18	31	Y	N
Project 19	6.25	7.20	19	7	Y	Y
Project 20	6.00	6.90	20	8	Y	N
Project 21	5.75	5.22	21	23	N	N
Project 22	5.00	6.80	22	10	N	N
Project 23	5.00	6.40	23	11	Y	N
Project 24	5.00	6.00	24	13	N	N
Project 25	5.00	5.40	25	20	N	N
Project 26	5.00	5.20	26	25	N	N
Project 27	5.00	4.67	27	28	N	N
Project 28	4.75	5.67	28	17	N	N
Project 29	4.67	5.50	29	19	N	N
Project 30	4.50	3.70	30	30	N	N
Project 31	4.00	7.22	31	6	N	Y

	Panel mean	Peer mean	Panel rank	Peer rank	Funded	Funded (peers)*
Project 32	3.75	2.56	32	32	N	N
Mean score:	6.370	5.798	Total funded:		20	7

*Assuming peers' judgement alone with a mean pass-mark of 7 were the sole criterion

** Bold indicates cases where panel and peer rank are more than ten places apart

Table 35: Comparing Panel and peer scores for the 2013/14 call

	Panel mean	Peer mean	Rank (panel)	Rank (peers)	Funded	Funded (peers)*
Project 1	8.8	6.6	1	12	Y	N
Project 2	8.3	8.3	2	2	Y	Y
Project 3	8.3	5.9	3	22	Y	N
Project 4	8.0	6.1	4	16	Y	N
Project 5	8.0	7.3	5	6	Y	Y
Project 6	7.8	7.3	6	7	Y	Y
Project 7	7.5	7.4	7	4	Y	Y
Project 8	7.3	8.8	8	1	Y	Y
Project 9	7.3	6.8	9	11	Y	N
Project 10	7.3	8.0	10	3	Y	Y
Project 11	7.0	6.0	11	20	Y	N
Project 12	7.0	7.4	12	5	Y	Y
Project 13	7.0	7.0	13	9	Y	Y
Project 14	6.7	5.7	14	23	N	N
Project 15	6.5	6.1	15	18	N	N
Project 16	6.3	6.1	16	17	N	N
Project 17	6.0	6.9	17	10	N	N
Project 18	5.8	5.5	18	25	N	N
Project 19	5.8	6.4	19	14	N	N
Project 20	5.8	7.3	20	8	N	Y
Project 21	5.7	6.3	21	15	N	N
Project 22	5.7	6.1	22	19	N	N
Project 23	5.5	6.5	23	13	N	N
Project 24	4.5	4.7	24	26	N	N
Project 25	4.0	6.0	25	21	N	N
Project 26	4.0	5.6	26	24	N	N
Mean score:	6.590	6.610	Total funded:		13	9

*Assuming peers' judgement alone with a mean pass-mark of 7 were the sole criterion

Table 36: Comparing Panel and peer scores for the 2014/15 call

	Panel mean	Peer mean	Rank (panel)	Rank (peers)	Funded	Funded (peers)*
Project 1	8.8	7.4	1	10	Y	Y
Project 2	8.4	8.3	2	2	Y	Y
Project 3	8.2	7.6	3	8	Y	Y
Project 4	8	7.3	4	11	Y	Y
Project 5	7.8	6.7	5	15	Y	N
Project 6	7.6	6.9	6	13	Y	N
Project 7	7.6	7.7	7	6	Y	Y
Project 8	7.6	9.1	8	1	Y	Y
Project 9	7.4	8	9	4	Y	Y
Project 10	7.2	7.5	10	9	Y	Y
Project 11	6.8	7.9	11	5	N	Y
Project 12	6.4	5.4	12	21	N	N
Project 13	6.4	8.3	13	3	Y	Y
Project 14	6.4	7.7	14	7	N	Y
Project 15	6.2	6.9	15	14	N	N
Project 16	6	6.5	16	16	Y	N
Project 17	5.7	6.5	17	17	N	N
Project 18	5.4	6	18	19	N	N
Project 19	5.2	6	19	20	N	N
Project 20	5	7.3	20	12	N	Y
Project 21	5	6.3	21	18	N	N
Project 22	4.6	5.3	22	22	N	N
Project 23	4	5.3	23	23	N	N
Project 24	3.8	4.9	24	24	N	N
Project 25	3.2	4.9	25	25	N	N
Mean score:	6.348	6.868	Total funded:		12	12

*Assuming peers' judgement alone with a mean pass-mark of 7 were the sole criterion

Appendix H Survey questions

Thank you for participating in this survey. Your views will help us to conduct a robust assessment of the ESRC's Transformative Research Scheme and enable conclusions to strengthen the funding landscape in the social sciences.

Completing the survey should take no more than ten minutes, and all information you provide will be treated in accordance to strict standards of research ethics and reported only in non-attributable form.

[click 'start' to begin the survey]

About you

Please provide the following information about yourself:

1. Gender
 - m/ f/ other
2. Age at the time of application
 - under 30
 - 30-39
 - 40-49
 - 50-59
 - over 60
3. Seniority at the time of application
 - research fellow (incl post-doc)
 - lecturer
 - senior lecturer
 - reader
 - professor
 - other academic position (specify)
 - I am not based at an academic research institution

About transformative research

4. Independently of the scheme, how would you describe 'transformative' research?

[box]

About your application

5. Please select the discipline most closely aligned with your application:
 - Area and Development Studies
 - Demography
 - Economics
 - Economic and Social History
 - Education

- Environmental Planning
- Human Geography
- Linguistics
- Management and Business Studies
- Political Science and International Studies
- Psychology
- Social Anthropology
- Social Policy
- Social Work
- Sociology
- Science and Technology Studies
- Socio-legal Studies
- Social Statistics, Methods and Computing
- Other [specify]

6. Please describe your project on the following criteria:

	Disagree Strongly	Disagree somewhat	Neutral	Agree Somewhat	Agree Strongly	Cannot judge
My project presented a major methodological innovation						
My project presented a major theoretical innovation						
My project represented a novel application of theory or methods to a new context						
My project was based on the engagement of unusual disciplinary and interdisciplinary perspectives						
My project involved an unusually high degree of risk in terms of likelihood to achieve successful research results						
My project was likely to involve an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants						
My project challenged core assumptions of the research establishment in my field						
This scope of my project was such that successful results could lead to a paradigm shift in its field						
My project was on a topic that is of high scientific interest to the research community						
The characteristics of my project pose challenges for eventual publication in leading (high impact factor) journals						
The nature of my project meant it would be unlikely to find favour with research council peer reviewers						
The results of my project (if successful) would have wide application outside of academia						
Achieving significant results from this research was likely to require long time frames and follow-up funding						

	Disagree Strongly	Disagree somewhat	Neutral	Agree Somewhat	Agree Strongly	Cannot judge
Even if successful, it will take a long time for results of this research to gain widespread acceptance in the academic community						

About your institution

7. Was there a selection process within your organisation that your application had to go through prior to submission to the ESRC?
 - Yes/no
8. Can you describe that process?
 - 1-stage
 - 2-stage
 - anonymous yes/ no /partially
 - peer or panel review
 - presentation/ interview
 - led by the department/ led by institution
9. How did you hear about the scheme?
 - My department or institution notified me about the scheme
 - My department or institution notified me about the scheme and recommended I should apply
 - I came across the scheme myself
 - The scheme was recommended to me by colleagues
 - Other [specify]

About the scheme

Review stage

10. The first stage of the application process is conducted anonymously. Do you feel that this focus on the project rather than the applicant(s) increased or decreased your chances of success?
 - Increased a lot
 - Increased somewhat
 - Neither
 - Decreased somewhat
 - Decreased a lot
11. Do you think anonymity of proposals is a measure that should be expanded to other research funding schemes?
 - No, and it should be removed from the ESRC Transformative research scheme
 - No, it works for the ESRC Transformative research scheme, but should not be used elsewhere.
 - Yes, in some other research funding schemes it would be helpful
 - Yes, anonymous proposal reviews should be used as widely as possible

12. Please share any further thoughts on the anonymity aspect of the application process

- [box]

13. Please indicate whether you agree or disagree with the following statements about the 2-page application format.

	Strongly agree	somewhat agree	neither	Somewhat disagree	Strongly disagree	Can't say
Two pages was too short to suitably describe the project						
The two-page format helped to crystallise the central idea(s) of the project						
The two page format meant it was easy to put the application together quickly						
The two page format made it attractive to apply to this scheme						

14. Were the assessment criteria for applications clear to you?

- Yes/ no

15. If 'no', please explain:

[box]

Pitch-to-peers

If your application was not shortlisted for the Pitch-to-peers workshop, please skip to the next section (question 21)

16. Please provide us with your views on the Pitch-to-peers workshop

	Strongly agree	somewhat agree	neither	Somewhat disagree	Strongly disagree	Can't say
The workshop was overall an enjoyable and rewarding exercise						
The workshop helped to identify transformative research proposals						
The workshop was well organised						
Input from assessors was helpful						
Input from peers was helpful						
Presenting in person was a good way for me to communicate my proposal						
Questions and discussion of my proposal focused on its transformative scope						
Questions and discussion of my proposal focused on its scholarly quality						
Questions and discussion of my proposal focused on its potential wider societal impact						
Questions and discussion of my proposal focused on its level of risk in terms of achieving successful outcomes						
Questions and discussion of my proposal focused on ethical challenges						

17. Please note any other comments you might have on the Pitch-to-peers workshop

- [box]

About the additional institutional award payment

Please answer the following questions only if you were awarded a grant in the 2012/2013 call to the Scheme. Otherwise skip to question 21.

18. In the 2012/13 call of the scheme, institutions with successful candidates received an additional payment of £50,000 for support of transformative research activities in the institution. Are you aware of this?

- Yes/ no

19. If yes, please describe briefly what this additional payment was used for

- [box]

20. have these activities funded through this additional payment in any way fostered a greater culture of transformative research at your institution?

- Yes/ no/ idk

Finally

21. Please feel free to note down any further comments you have about the ESRC's Transformative research scheme:

[box]

Thank you for participating in this survey! The information you have submitted will be treated in accordance with high standards of research ethics and will only be reported in non-attributable form. Should you have any further questions about the study, you can contact the project manager, Dr Peter Kolarz (peter.kolarz@technopolis-group.com).

Appendix I Phase 2: Interview details

Interviews were conducted via telephone from March until May 2016. The interviews took a semi-structured approach, but the main questions asked to all interviewees were:

- What are the main outputs from your TR project?
- Do you think your TR project was successful?
- Has your project been transformative? In what way?
- Was it at all necessary to deviate from your original project plan? If so, can you explain what caused the deviation(s)?
- Can you comment on the wider significance of your project? Have any other academics or stakeholder groups taken notice?
- Next steps: what are your plans?
- Have there been any difficulties caused by the transformative nature of the work?
- How satisfied are you with your relationship with ESRC over the grant period, including reporting, support, etc.?
- Having completed a TR grant, can you comment on how you see this scheme?
- What would have happened if your proposal had not been accepted?

Table 37: List of Phase 2 interviewees

Inter-viewee	Call year	Project title	Institution	Aligned discipline	Inter-viewer
Dr Alice Street	2012-2013	'Off the Grid': relational infrastructures for fragile futures	Edinburgh University	Social Anthropology	Stephanie Gardham
Prof Alison Liebling	2012-2013	Locating trust in a climate of fear: religion, moral status, prisoner leadership, and risk in maximum security prisons	Cambridge University	Socio-legal studies	Stephanie Gardham
Prof George Davey Smith	2012-2013	The biosocial archive: transforming lifecourse social research through the incorporation of epigenetic measures	Bristol University	Sociology	Martin Wain
Prof Hilary Graham	2012-2013	Health of Populations and Ecosystems (HOPE)	York University	Human geography	Peter Kolarz
Prof Mark Whitehead	2012-2013	Negotiating neoliberalism: changing behaviours, values and beliefs	Aberystwyth University	Human geography	Stephanie Gardham
Prof Nicky Gregson	2012-2013	Illicit economics and the spaces of circulation	Durham University	Human geography	Stephanie Gardham
Prof Nikolas Rose	2012-2013	A new sociology for a new century: transforming the relations between sociology and neuroscience, through a study of mental life and the city	King's College London	Sociology	Stephanie Gardham
Dr Paul Warren	2012-2013	A new perspective on human judgement and decision making as optimal: A framework for behaviour change	Manchester University	Psychology	Stephanie Gardham
Prof Ran Spiegler	2012-2013	Games between diversely sophisticated players	UCL	Economics	Stephanie Gardham
Prof Stephen Hinchliffe	2012-2013	Contagion: transforming social analysis and method	Exeter University	Human geography	Stephanie Gardham
Prof Stephen Reicher	2012-2013	Beyond the 'banality of evil': a new understanding of conformity and atrocity	St Andrews University	Psychology	Stephanie Gardham

Inter-viewee	Call year	Project title	Institution	Aligned discipline	Inter-viewer
Dr Vincent Reid	2012-2013	Understanding light in the late term human fetus: Proof of concept for social research techniques	Lancaster University	Psychology	Stephanie Gardham
Dr Farida Vis	2013-2014	Picturing the social: transforming our understanding of images in social media and Big Data research	Sheffield University	Sociology	Stephanie Gardham
Dr Guido Orgs	2013-2014	Synchronous movement cooperation and the performing arts	Brunel University	Psychology	Peter Kolarz
Prof Jane Raymond	2013-2014	Understanding cognition in middle adulthood	Birmingham University	Psychology	Stephanie Gardham
Prof John Garry	2013-2014	Randomly selected "politicians": Transforming democracy in the post-conflict context	Queen's University Belfast	Political Science and International Studies	Peter Kolarz
Dr Jonathan Halket	2013-2014	Urban dynamics in a complex world: The spatial dynamics of housing	Institute of Fiscal Studies	Economics	Stephanie Gardham
Dr Lauren Devine	2013-2014	Rethinking child protection strategy: evaluating research findings and numeric data to challenge whether current intervention strategy is justified	UWE	Socio-legal studies	Stephanie Gardham
Prof Petroc Sumner	2013-2014	Understanding and optimising health-related press releases as complex public health interventions	Cardiff University	Psychology	Stephanie Gardham
Dr Rebecca Marsland	2013-2014	Beelines	Edinburgh University	Social Anthropology	Peter Kolarz

Table 38: Summary of interviews

Interview question	Response summary (all numbers are out of 20)
What are the main outputs from your TR project?	<ul style="list-style-type: none"> • 12 say there are papers in progress or under review • 10 noted outreach/ public engagement activities among main outputs • 5 have apps/ videos/ films etc. (unusual outputs)
Do you think your TR project was successful?	<ul style="list-style-type: none"> • 7 say enthusiastically yes, outright • 10 say yes, but more to do • 5 a little more cautious • 1-2 fairly despondent
Has your project been transformative? In what way?	<ul style="list-style-type: none"> • 16 say 'yes' in some form • 8 of these say yes, but it is 'early days' and in some ways too soon to tell • 4 tend towards 'no'
Was it at all necessary to deviate from your original project plan?	<ul style="list-style-type: none"> • 6 say 'no' • 8 say 'yes', but only slight • 2 said severe changes
Can you comment on the wider significance of your project? Have any other academics or stakeholder groups taken notice?	<ul style="list-style-type: none"> • 3 had highly significant public/ policy interest • 11 had some, including some policy interest • 6 had none or very low non-academic interest

Interview question	Response summary (all numbers are out of 20)
Next steps: what are your plans?	<ul style="list-style-type: none"> • 6 have secured further funding • 8 have grants applications in progress or under review • Others have been unsuccessful or do not have clear plans yet.
Have there been any difficulties caused by the transformative nature of the work?	<ul style="list-style-type: none"> • Very mixed. No major significant issues. Some publishing-related, but usually due to interdisciplinary problems.
How satisfied are you with your relationship with ESRC over the grant period, including reporting, support, etc.?	<ul style="list-style-type: none"> • 7 (confirmed) extensions granted, most applied for one! • 4 critical comments on ResearchFish • Only one who had problems/ was unhappy
Having completed a TR grant, can you comment on how you see this scheme?	<ul style="list-style-type: none"> • 18 say it's a great/ fantastic/ awesome scheme overall. • 10 say 18 months is too short • 8 say there should be some form of follow-on funding • 4 say subsequent 'cohort' meetings would be good • 2 say have a tiered or halfway review approach
What would have happened if your proposal had not been accepted?	<ul style="list-style-type: none"> • 5 say outright the project would not have happened • 5 say unlikely to have happened w/o TR call • 5 say would have gone ahead but slower or with reduced scope • 3 said they would have tried elsewhere • 2 unsure/ unclear

Appendix J Phase 2: Details of Phase 2 review exercise

J.1 Method

Our peer review exercise of outputs ran from the 18th of March to the 29th of April 2016. WE used the same pool of reviewers as we did for the review of applications in Phase 1 (with the exception of Rupert Brown, who was unavailable for this second exercise). All reviewers had prior experience as RAE/REF panellists.

We selected a total of 80 outputs (journal articles and working papers only) logged in ResearchFish: 40 from the TR scheme and 40 from the ESRC standard grants scheme. To ensure a closest possible comparison, we limited the Standard grants to those of a value of between £180k and £320k, with start dates and end dates likewise approximating as closely as possible to those of the TR grants. Each output was assigned to two reviewers, based on their stated areas of expertise and out internal assessment of the thematic content of the outputs. Reviewers were not informed either about which outputs came from which scheme, or that the review involved outputs from Standard grants at all. However, as some articles acknowledge the relevant funding scheme, some level of disclosure was present.

There was no sampling process as such: we imposed the additional limitation of selecting no more than three outputs from each project – within projects with more than three listed outputs, we selected at random. With this in place, it was only just possible to compile 40 outputs from the two grant schemes. Where a very small number of excess outputs was still available, we eliminated a small number at random to bring the total down to 40 in each scheme. In this sense, the exercise covers a near-exhaustive share of outputs logged on ResearchFish for the TR scheme, as well as for standard grants of a comparable size and timeline.

For each output, reviewers were asked to complete the template shown in the sub-section below. Once returned, we analysed at individual review, output and project level. Judgements were converted to numerical values (so 'disagree strongly' or 'very low' is equivalent to '1', agree strongly' or 'very high' equates to '5').

Critically, we also asked reviewers to rate, for each review, their own confidence in their judgements on a four-point scale. In our analysis we consistently found that patterns and tendencies emerged more clearly when we excluded reviews whose confidence level was self-assessed as 'less confident' or 'not at all confident'

J.2 Review template

ESRC Transformative Research

Outcomes review, Spring 2016

Output short title:	
----------------------------	--

Reviewer name:	
-----------------------	--

Please rate the output on each of the following criteria:

	Disagree Strongly	Disagree somewhat	Neutral	Agree Somewhat	Agree Strongly	Cannot judge
This output presents a pioneering theoretical or methodological innovation						
This output represents a novel application of theory or methods to a new context						
This output is based on the engagement of unusual interdisciplinary perspectives						
This research for this output likely involved an unusually high degree of risk in terms of likelihood to achieve successful research results						
This output likely involved an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants						
This research behind this output may generally be described as 'high-risk, high reward'						
This output is likely to challenge widespread assumptions of typical social scientific enquiry						
This output could lead to a paradigm shift in its field						
This output is likely to raise considerable interest in the social scientific community						
This output lends itself to eventual application of results (if successful) outside of academia						
This output is likely to require comparatively long lead times to achieve successful results						
This output is more typical of early career researchers than of senior, established scholars						
This output may struggle to gain widespread acceptance in the academic community						
This output is likely to produce a broad base of knowledge, new thinking or insights						
The research behind this output is likely to require comparatively high amounts of follow-up funding to achieve genuinely salient results						

Please also rate the quality of the output on the following criteria. We take the former three from the REF2014 criteria and add scientific and societal transformative scope as a fourth and fifth.

	None/ very low	Low	Moderate	High	Very high	Cannot judge
Originality						
Significance						
Rigour						
Scope to transform social science						
Scope to transform wider society						

Overall, would you describe this project as ‘transformative’?

Yes	
No	
Can't say	

Please explain your answer (briefly):

Please add any further thoughts, qualifications or comments you feel might be important:

Finally, please give an indication of how confidently you feel you have been able to rate this output:

Very confident – this output is well within my area(s) of expertise	
Fairly confident – the output is not fully within my area(s) of expertise, but I could still make a fair judgement	
Less confident – I could provide some level of judgement, but I have limited expertise and am unsure about several elements	
Not confident – this output is in large parts beyond my expertise	

J.3 Results

Table 39: Review exercise – mean and median figures across the criteria (TR/ standard split)

	All reviews (160)						Only reviews self-judged as 'very' or 'fairly confident'					
	TR Scheme (80)		Standard scheme (80)		Difference		TR Scheme (47)		Standard scheme (68)		Difference	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
This output presents a pioneering theoretical or methodological innovation	3.34	4	3.54	4	-0.20	0	3.6	4	3.5	4	0.1	0
This output represents a novel application of theory or methods to a new context	3.33	4	3.67	4	-0.35	0	3.7	4	3.7	4	0.0	0
This output is based on the engagement of unusual interdisciplinary perspectives	2.68	2	2.93	3	-0.24	-1	3.3	4	2.8	2.5	0.5	1.5
This research for this output likely involved an unusually high degree of risk in terms of likelihood to achieve successful research results	2.18	2	2.00	2	0.18	0	2.3	2	2.0	2	0.3	0
This output likely involved an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants	1.37	1	1.24	1	0.13	0	1.5	1	1.4	1	0.1	0
This research behind this output may generally be described as 'high-risk, high reward'	2.07	2	1.96	2	0.11	0	2.2	2	2.1	2	0.1	0
This output is likely to challenge widespread assumptions of typical social scientific enquiry	3.22	4	2.94	3	0.28	1	3.4	4	2.9	3	0.5	1
This output could lead to a paradigm shift in its field	3.26	3.5	3.12	3	0.14	0.5	3.4	4	3.1	3	0.3	1
This output is likely to raise considerable interest in the social scientific community	3.53	4	3.91	4	-0.38	0	3.9	4	4.0	4	-0.1	0
This output lends itself to eventual application of results (if successful) outside of academia	3.98	4	4.05	4	-0.07	0	4.2	4	4.1	4	0.0	0

	All reviews (160)						Only reviews self-judged as 'very' or 'fairly confident'					
	TR Scheme (80)		Standard scheme (80)		Difference		TR Scheme (47)		Standard scheme (68)		Difference	
	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median	Mean	Median
This output is likely to require comparatively long lead times to achieve successful results	2.89	3	2.97	3	-0.08	0	2.9	3	2.9	3	0.0	0
This output is more typical of early career researchers than of senior, established scholars	2.33	2	2.49	2	-0.16	0	2.2	2	2.5	2	-0.3	0
This output may struggle to gain widespread acceptance in the academic community	2.58	2	2.44	2	0.14	0	2.5	2	2.3	2	0.2	0
This output is likely to produce a broad base of knowledge, new thinking or insights	3.47	4	3.75	4	-0.28	0	3.7	4	3.8	4	0.0	0
The research behind this output is likely to require comparatively high amounts of follow-up funding to achieve genuinely salient results	2.91	3	2.64	2	0.28	1	3.0	3	2.6	2	0.4	1
ORIGINALITY	3.35	3	3.51	3	-0.16	0	3.6	4	3.5	3	0.1	1
SIGNIFICANCE	3.54	4	3.78	4	-0.24	0	3.8	4	3.8	4	0.1	0
RIGOUR	3.57	4	3.79	4	-0.23	0	3.8	4	3.8	4	0.1	0
TR SCOPE - Social Science	2.98	3	3.13	3	-0.15	0	3.4	3	3.1	3	0.2	0
TR SCOPE - Society	3.32	3	3.39	3	-0.07	0	3.3	3	3.4	3	-0.1	0

Table 40: Overall verdicts

'Overall, would you describe this output as 'transformative?'	TR scheme (47)		Standard grants (68)	
	Count	%	Count	%
Yes	28	60%	37	54%
No	15	32%	24	35%
Cannot say	4	9%	7	10%

Table 41: % of reviews that received the highest possible mark

	TR Scheme - % of reviews with the top mark (/5)	Standard grants - % of reviews with the top mark (/5)	Difference
This output presents a pioneering theoretical or methodological innovation	30%	13%	17%
This output represents a novel application of theory or methods to a new context	30%	18%	12%
This output is based on the engagement of unusual interdisciplinary perspectives	19%	6%	13%
This research for this output likely involved an unusually high degree of risk in terms of likelihood to achieve successful research results	2%	0%	2%
This output likely involved an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants	2%	0%	2%
This research behind this output may generally be described as 'high-risk, high reward'	2%	0%	2%
This output is likely to challenge widespread assumptions of typical social scientific enquiry	17%	4%	13%
This output could lead to a paradigm shift in its field	17%	12%	5%
This output is likely to raise considerable interest in the social scientific community	23%	21%	3%
This output lends itself to eventual application of results (if successful) outside of academia	36%	32%	4%
This output is likely to require comparatively long lead times to achieve successful results	4%	4%	0%
This output is more typical of early career researchers than of senior, established scholars	0%	7%	-7%
This output may struggle to gain widespread acceptance in the academic community	4%	1%	3%
This output is likely to produce a broad base of knowledge, new thinking or insights	21%	19%	2%
The research behind this output is likely to require comparatively high amounts of follow-up funding to achieve genuinely salient results	13%	3%	10%
ORIGINALITY	17%	10%	7%
SIGNIFICANCE	23%	24%	0%
RIGOUR	28%	13%	14%
TR SCOPE - Social Science	15%	6%	9%
TR SCOPE - Society	6%	9%	-2%

NB: excludes reviews where reviewers did not rate themselves at least 'fairly' or 'very confident'

Table 42: Differences between reviews for 2012-13 and 2013-14 TR outputs

	2012-13 call (29)		2013-14 call (16)		Differences	
	Mean	Median	Mean	Median	Diff. mean	Diff. median
This output presents a pioneering theoretical or methodological innovation	3.97	4.00	2.94	2.50	1.03	1.50
This output represents a novel application of theory or methods to a new context	4.20	4.00	2.88	2.00	1.33	2.00
This output is based on the engagement of unusual interdisciplinary perspectives	3.81	4.00	2.31	2.00	1.49	2.00
This research for this output likely involved an unusually high degree of risk in terms of likelihood to achieve successful research results	2.37	2.00	2.20	2.00	0.17	0.00
This output likely involved an unusually high degree of risk in terms of the safety and wellbeing of the researcher and/ or participants	1.68	1.00	1.19	1.00	0.49	0.00
This research behind this output may generally be described as 'high-risk, high reward'	2.37	2.00	1.81	2.00	0.55	0.00
This output is likely to challenge widespread assumptions of typical social scientific enquiry	3.71	4.00	2.94	3.50	0.77	0.50
This output could lead to a paradigm shift in its field	3.83	4.00	2.63	2.00	1.21	2.00
This output is likely to raise considerable interest in the social scientific community	4.23	4.00	3.25	3.00	0.98	1.00
This output lends itself to eventual application of results (if successful) outside of academia	4.34	4.00	3.81	4.00	0.53	0.00
This output is likely to require comparatively long lead times to achieve successful results	3.17	3.50	2.50	2.00	0.67	1.50
This output is more typical of early career researchers than of senior, established scholars	1.83	2.00	2.81	3.00	-0.98	-1.00
This output may struggle to gain widespread acceptance in the academic community	2.42	2.00	2.69	2.50	-0.27	-0.50
This output is likely to produce a broad base of knowledge, new thinking or insights	4.03	4.00	3.19	3.00	0.84	1.00
The research behind this output is likely to require comparatively high amounts of follow-up funding to achieve genuinely salient results	3.20	3.00	2.63	2.00	0.58	1.00
ORIGINALITY	3.74	4.00	3.25	3.00	0.49	1.00
SIGNIFICANCE	4.10	4.00	3.33	3.00	0.76	1.00
RIGOUR	3.94	4.00	3.63	3.50	0.31	0.50
TR SCOPE - Social Science	3.65	3.00	2.88	3.00	0.77	0.00
TR SCOPE - Society	3.48	3.00	3.07	3.00	0.42	0.00

technopolis |group| United Kingdom
3 Pavilion Buildings
Brighton BN1 1EE
United Kingdom
T +44 1273 204320
E info@technopolis-group.com
www.technopolis-group.com