



The British
Psychological Society
Research Board

Psychology Research Excellence Framework 2014 Impact Analysis

Thom Baguley

Acknowledgement

The Research Board sincerely thanks Professor Thomas Baguley for preparing this report on its behalf.

Thom Baguley would like to thank Harriet Smith, Natalie Harrison and Derek Moore for their invaluable assistance in developing the analyses and in preparing this report.

© The British Psychological Society 2017

The British Psychological Society

St Andrews House, 48 Princess Road East, Leicester LE1 7DR, UK

Telephone 0116 254 9568 Facsimile 0116 247 0787

E-mail mail@bps.org.uk Website www.bps.org.uk

Incorporated by Royal Charter Registered Charity No 229642

If you have problems reading this document and would like it in a different format, please contact us with your specific requirements.

Tel: 0116 252 9523; E-mail: P4P@bps.org.uk.

Contents

1. Executive summary	2
2. Background	3
2.1 The data set	3
2.1.1 The case studies	3
2.1.2 HEFCE summary impact type	4
2.1.3 Research Excellence Framework 2014 data	4
2.1.4 Coding case studies by core or applied area of psychology	4
2.1.5 Statistical modelling	6
2.2 Summary of results	7
2.2.1 HEFCE summary impact type	7
2.2.2 Simultaneous analysis of all areas	7
2.2.3 Analysis of core psychology areas	10
2.2.4 Analysis of applied psychology areas	10
2.2.5 Conclusions	12

1. Executive summary

- The report summarises the outcomes of an analysis of published Research Excellence Framework data on 315 impact case studies submitted to Unit of Assessment (UoA) A04: Psychology, Psychiatry and Neuroscience.
- Analysis of the 81 submissions made reveals that on average each submission included 4.9 case studies. However, there is significant variation between submissions, with the very largest including 30 case studies.
- Of the categories of impact type used by Higher Education Funding Council for England (HEFCE) (cultural, economic, environmental, health, legal, political, societal and technological), health and societal dominate the case studies submitted to UoA A04. However, this has caused concern in the psychological research community that the HEFCE classifications are arguably superficial and do not necessarily reflect the actual type of impact achieved (i.e. based on the topic of the research rather than the impact itself).
- Re-analysis of the case studies focused on recoding them into core areas (biological, cognitive, social and developmental psychology) and applied (clinical, counselling, educational, forensic, health, neuropsychology, occupational, and sport and exercise).
- Many case studies were identified as interdisciplinary (combining multiple core and applied areas).
- The most commonly coded areas were identified as cognitive, health and social (with cognitive and social each associated with nearly 40 per cent of case studies). Less commonly coded areas were counselling, neuropsychology, occupational, and sport and exercise.
- This reflects concerns that sports and exercise psychology and occupational/organisational psychology was returned to other non-psychology sub-panels.
- However, similar concerns regarding other areas, such as social psychology, do not seem to extend to the impact case studies.
- Regression analysis on the raw impact data for the core areas of psychology reveals that case studies drawing on two or more core areas have a higher predicted impact.
- However, once adjustment for size and strength of submission is made, all core areas have a similarly high level of predicted impact. There is no evidence, for instance, that social psychology underperforms.
- Analysis for the applied areas shows that case studies that draw on multiple areas (health psychology and forensic psychology) have higher predicted impact. Predicted impact is lower for other applied areas (notably sport and exercise and neuropsychology) after controlling for size and strength of submissions. However, this observation rests on a small number of case studies and should be interpreted with caution.
- In conclusion, the impact of psychological research submitted to Research Excellence Framework 2014 draws widely from the core areas of the discipline and most applied areas. In addition, there is little evidence that impact grades vary according to the areas of psychology that a case study draws on.

2. Background

Since Research Excellence Framework 2014 there has been considerable interest in how different areas of the psychology discipline contributed to and were represented in submissions from psychology departments. However, much of the discussion around these topics has been speculative. The project reported aims to provide preliminary evidence that can inform further discussion (and perhaps further analysis).

The analysis draws primarily on publicly Research Excellence Framework data in order to explore a range of questions, including:

- What areas of psychology do impact case studies draw on?
- Is the spread of case studies even across core or applied areas?
- Does the core or applied area predict on the impact grades?

2.1 The data set

The analysis draws on published Research Excellence Framework data on impact case studies in UoA A04 Psychology, Psychiatry and Neuroscience. In total there are 315 case studies in A04, arising from 81 submissions (excluding one submission with no case studies).

2.1.1 The case studies

On average there are 4.9 case studies per submission, with a median of 2. Of the 81 submissions that included case studies the size of return, and hence number of case studies is highly skewed – with the largest submission including 30 case studies (see Figure 1).

Figure 1: Spread of coding categories across case studies

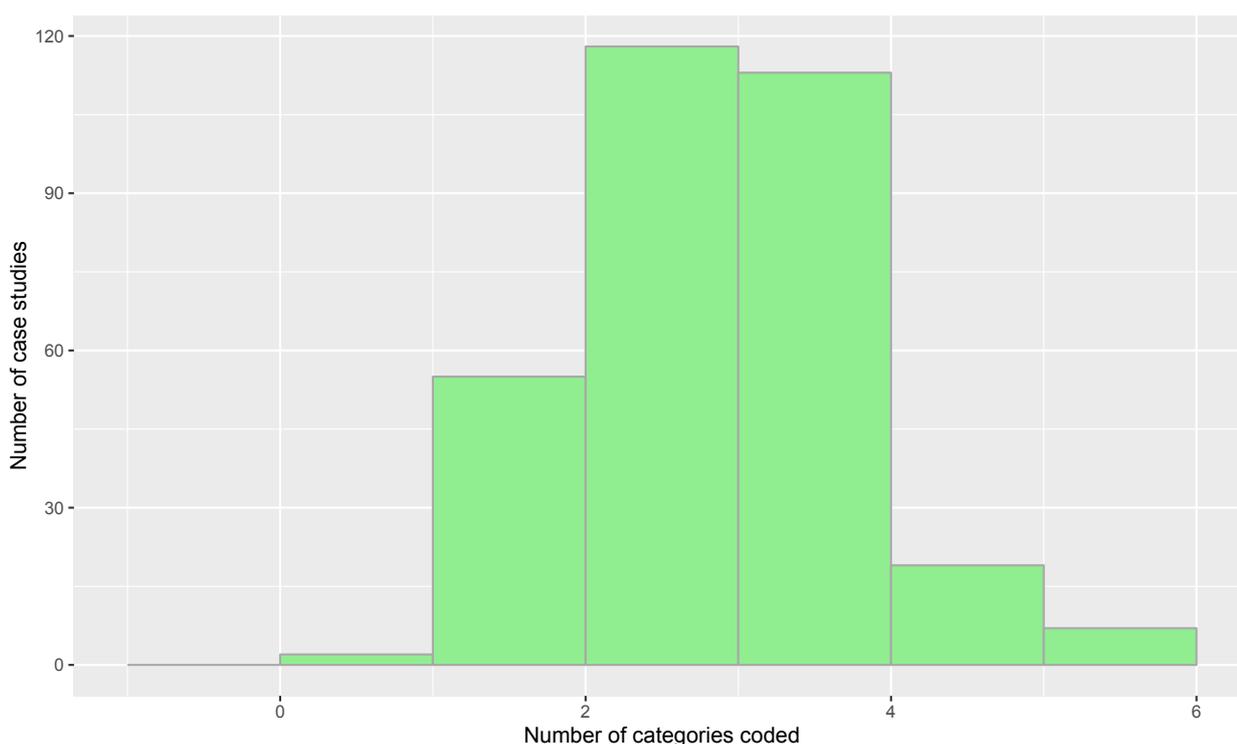
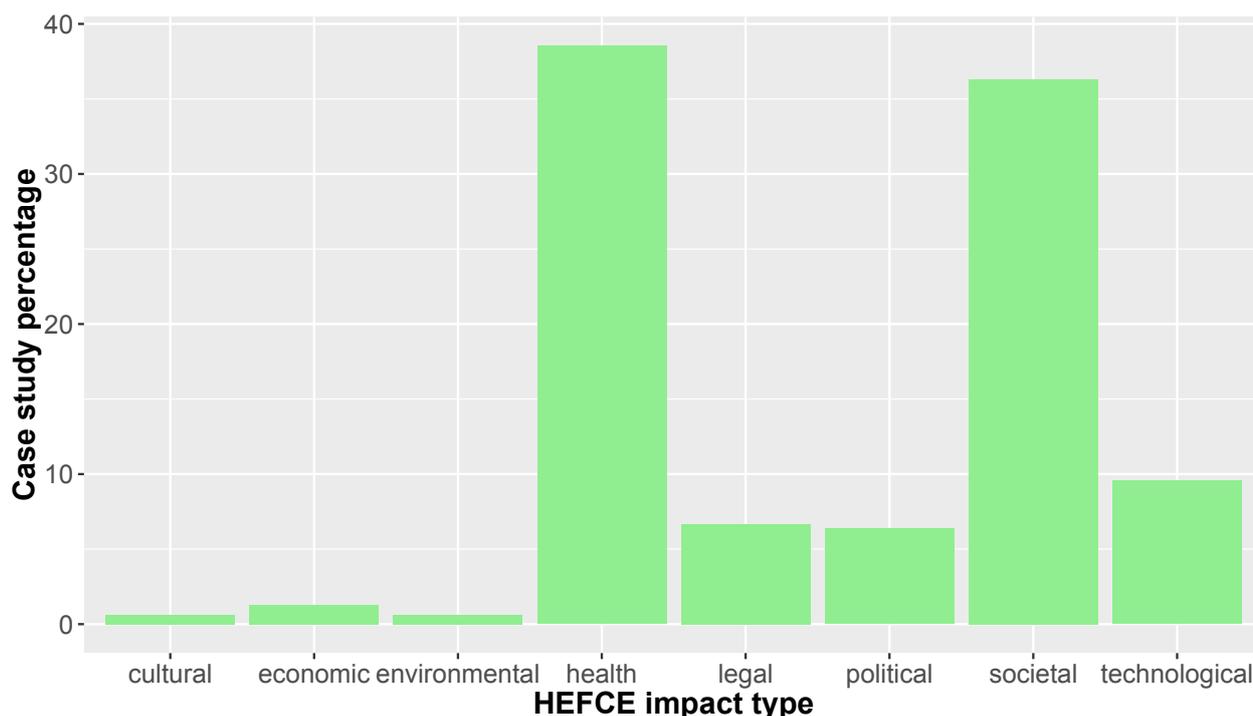


Figure 2: Percentage of case studies by HEFCE impact type



2.1.2 HEFCE summary impact type

Each case study is categorised by Higher Education Funding Council for England (HEFCE) under one of eight summary impact types: cultural, economic, environmental, health, legal, political, societal and technological. For UoA A04 two impact types dominate: health and societal. Three are hardly represented at all: cultural, economic and environmental (see Figure 2). This is a little puzzling as many of the case studies do have economic impacts and (anecdotally at least) the classifications are sometimes superficial. For example, forensic psychology case studies tend to be classified as legal even when the impact seems to be societal or technological.

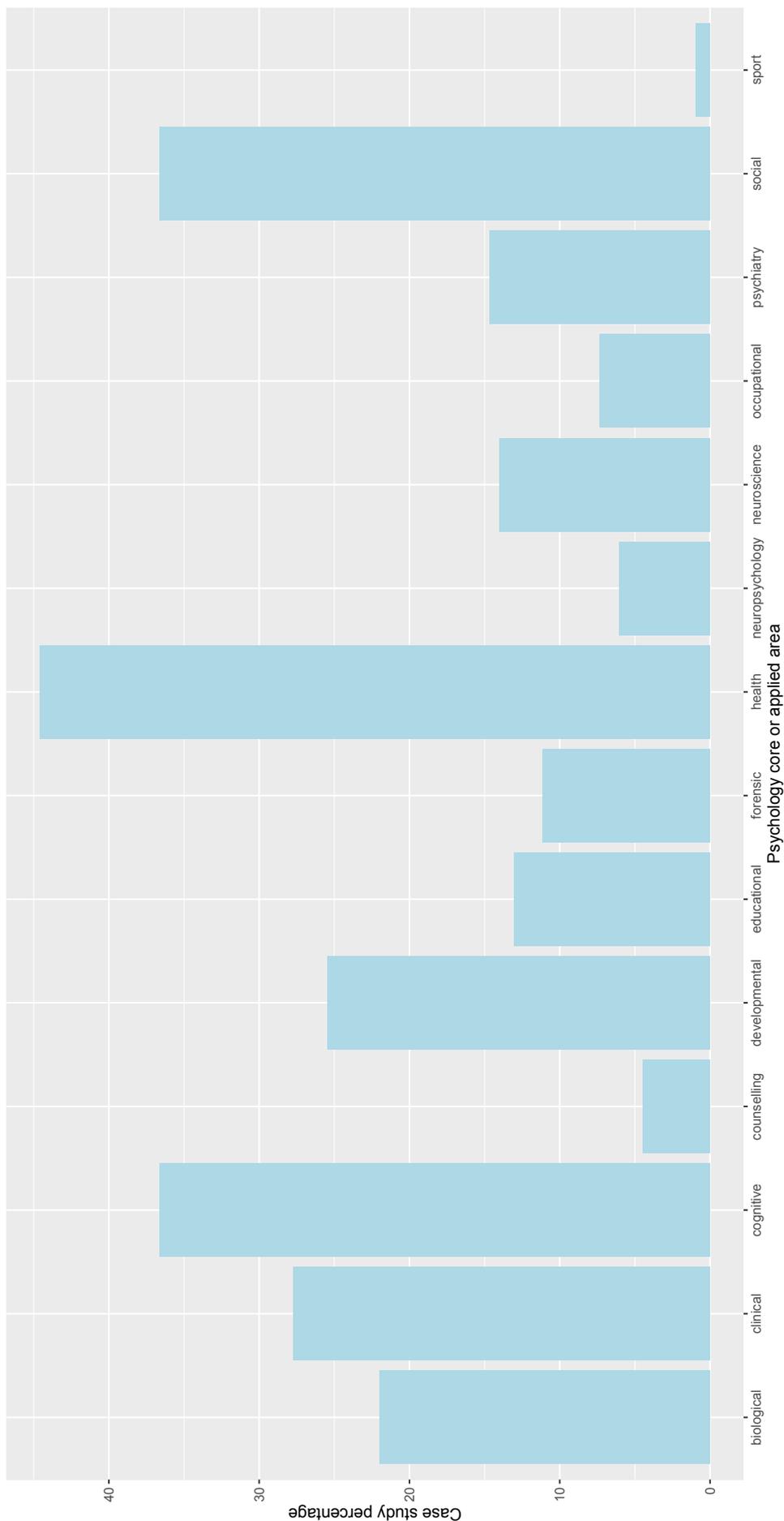
2.1.3 Research Excellence Framework 2014 data

While some Research Excellence Framework data is not published (e.g. grades assigned to individual case studies) data for outputs, environment and impact are available aggregated at submission level in addition to some contextual information such as FTE returned. These analyses here use the grade point average (GPA) for outputs, environment and impact. This produces a metric ranging from 0 to 4 (ranging from 100 per cent unclassified to 100 per cent 4★ – though in practice scores below 2 are rare). The FTE metric captures both the number of staff returned cultural economic environmental health legal political societal technological HEFCE impact type Case study percentage and the number of case studies (itself a function of the FTE). Owing to the highly skewed distribution of FTEs this variable was log transformed when included in any analysis.

2.1.4 Coding case studies by core or applied area of psychology

Each case study was coded independently by two psychology research assistants (NH and HS) into core and applied areas of psychology. Core areas were defined using the Quality Assurance Agency (QAA) benchmark categories: biological, cognitive, developmental,

Figure 3: Percentage of case studies by core or applied area



social (but excluding individual differences). The applied areas were based on the Society's divisions (excluding Teachers and Researchers in Psychology): Clinical, Counselling, Educational, Forensic, Health, Neuropsychology, Occupational, Sport and Exercise. Early in the coding it was clear that many case studies were interdisciplinary (in the sense that they combined multiple core and multiple applied areas). After all case studies were coded the two researchers met to agree the final coding – though in most cases there was close agreement. The coders were unable to agree a primary core or applied area for each case study. In addition, two other codes, psychiatry and neuroscience, were used to capture work from other disciplines (however, in practice many case studies drew on psychology and one or more other disciplines).

Figure 3 shows the percentage of case studies coded within each area. Cognitive, health and social are the most common with counselling, neuropsychology, occupational and sport and exercise less common. The lack of case studies codes as sport and exercise, or occupation is line with suggestions that psychologists working in these areas are more likely to be returned to other units. However, similar concerns around areas such social psychology seem not to extend to impact case studies.

2.1.5 Statistical modelling

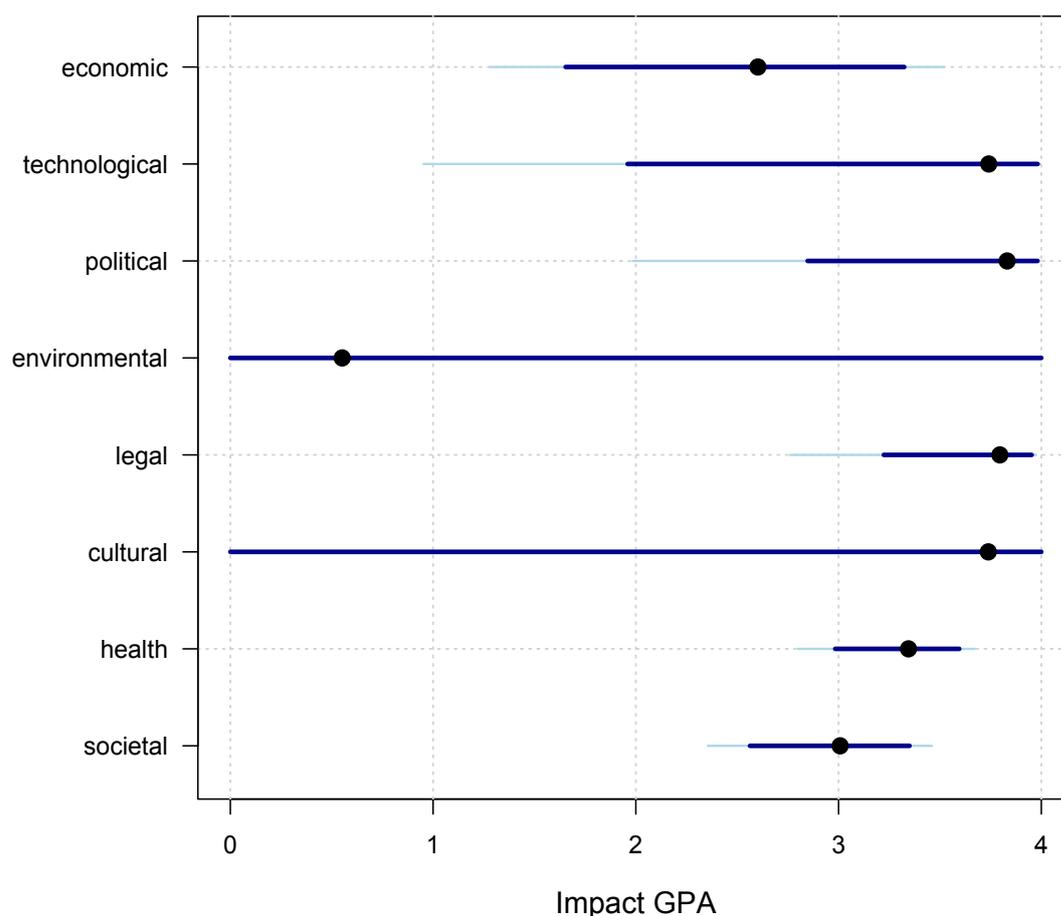
The main analyses involve predicting impact GPA from case study characteristics. Finding an appropriate model presents several challenges. First, impact GPA is bounded at 0 and 4. This is particularly problematic because the impact GPA is negatively skewed – with a lot of high scoring case studies and a long tail of lower scoring ones. Second, the case studies are clustered within submissions (hierarchical or multilevel model). Third, this is case of a micro–macro multilevel model in which a level 2 outcome (impact GPA for a submission) are predicted by level 1 data (case study characteristics).

The first problem can be overcome by using an empirical logistic transformation. After rescaling impact GPA to the range 0 to 1 the data are transformed to a logit scale (with a small constant added to numerator and denominator in the transformation to handle 0 or 1 outcomes). Parameter estimates can then be transformed back to the 0 to 4 scale after analysis.

The second problem would normally be handled using a multilevel model, but conventional multilevel models are inaccurate for micro to macro prediction. Various solutions exist, but all have limitations. However, Kromrey & Johnson-Fisher (2015) have presented data suggesting that a relatively simple approach of using level 1 averages (aggregated over level 2 units) in conjunction with a correction to handle heterogeneity is relatively unbiased and high in statistical power. This is the approach adopted here – which has the added virtue of being relatively simple to set up in R (an open source statistical software environment).

For some analyses covariates such as output GPA, environment GPA and $\log(\text{FTE})$ were included in the model (after centering to make it easier to interpret the other effects). Together these three are strongly predictive of impact GPA (though they are correlated with each other and only output GPA predicts impact GPA uniquely). In models with these covariates added, the predictions are adjusted to estimate a submission with mean output GPA, mean environment GPA and median FTE (the mean \log FTE corresponding to median FTE).

Figure 4: Predicting impact GPA by HEFCE impact type



2.2 Summary of results

Brief graphical summaries of the results are presented below. In all cases the predicted mean impact GPA is shown with error bars. Two types of intervals are depicted using two-tiered error bars. The outer tier (thin, light blue lines) depict conventional 95 per cent confidence intervals for the predicted mean. These give an indication of the precision of the predicted mean (a function of effective sample size and variability in the data). The outer tier (thick, dark blue lines) are adjusted so that intervals that overlap correspond (approximately) to statistically significant differences with $\alpha = .05$ (see Baguley, 2012). Note that no corrections for multiple testing have been applied.

2.2.1 HEFCE Summary impact type

Figure 4 shows predicted impact GPA by summary impact type. The most striking pattern is for environmental and cultural impact where there are simply too few data points to estimate anything reliably. For the other summary impact types, the predicted averages tend to be high and no clear patterns of difference detectable.

2.2.2 Simultaneous analysis of all areas

Figure 5 shows the predicted mean impact GPA for a simultaneous analysis of all coded areas. It is worth noting that these predicted means are for the unique contribution of each area. The main patterns that emerge are the consistently strong predicted impact for the biological area (close to ceiling) and a number of areas (cognitive, developmental, social, clinical, counselling, educational, forensic and health) with predicted impact slightly below that. Occupational, neuropsychology and psychiatry also do well but with

Figure 5: Predicting impact GPA by core or applied area

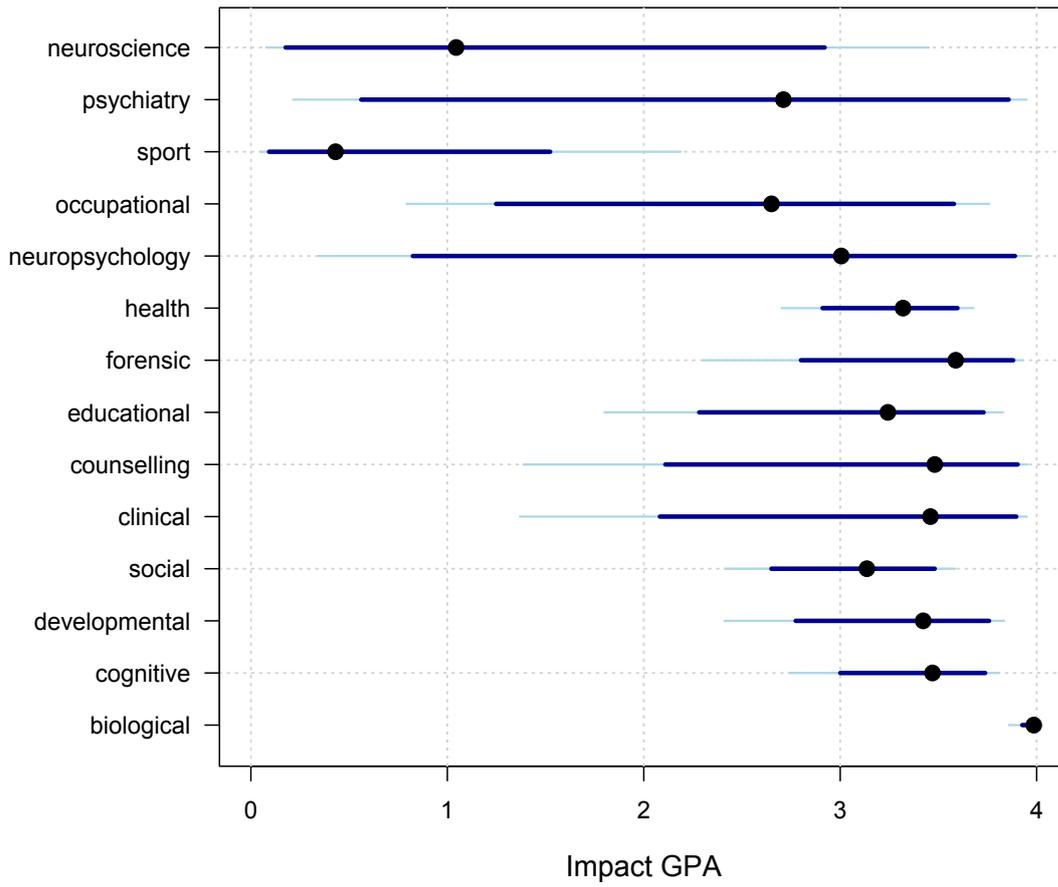


Figure 6: ...adjusting for output GPA, environment GPA and log(FTE)

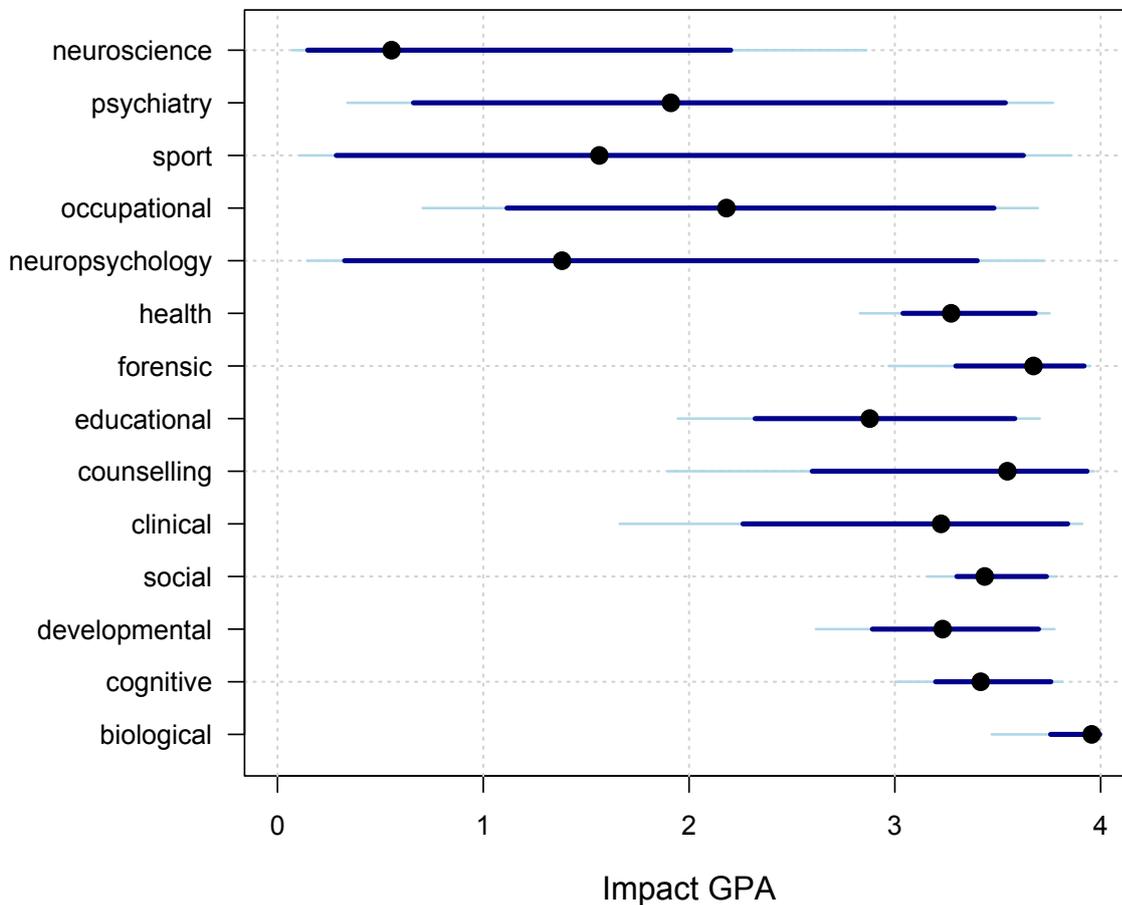


Figure 7: Predicting impact GPA by core area

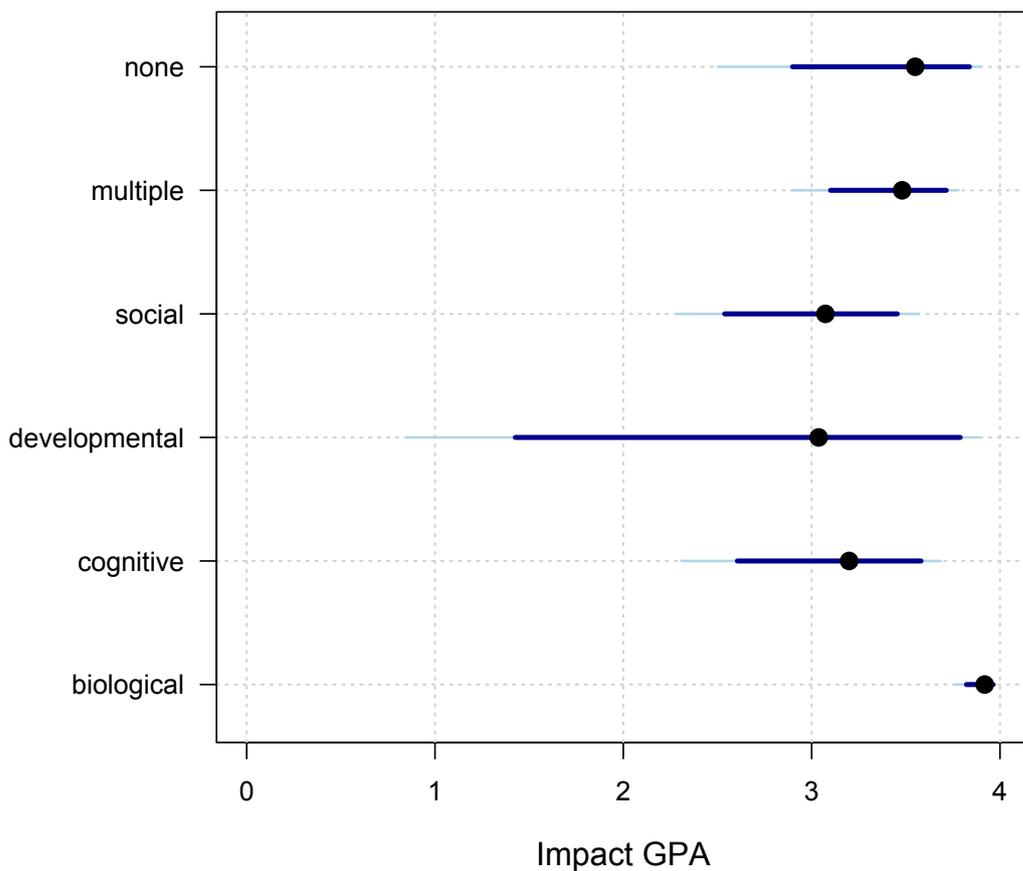
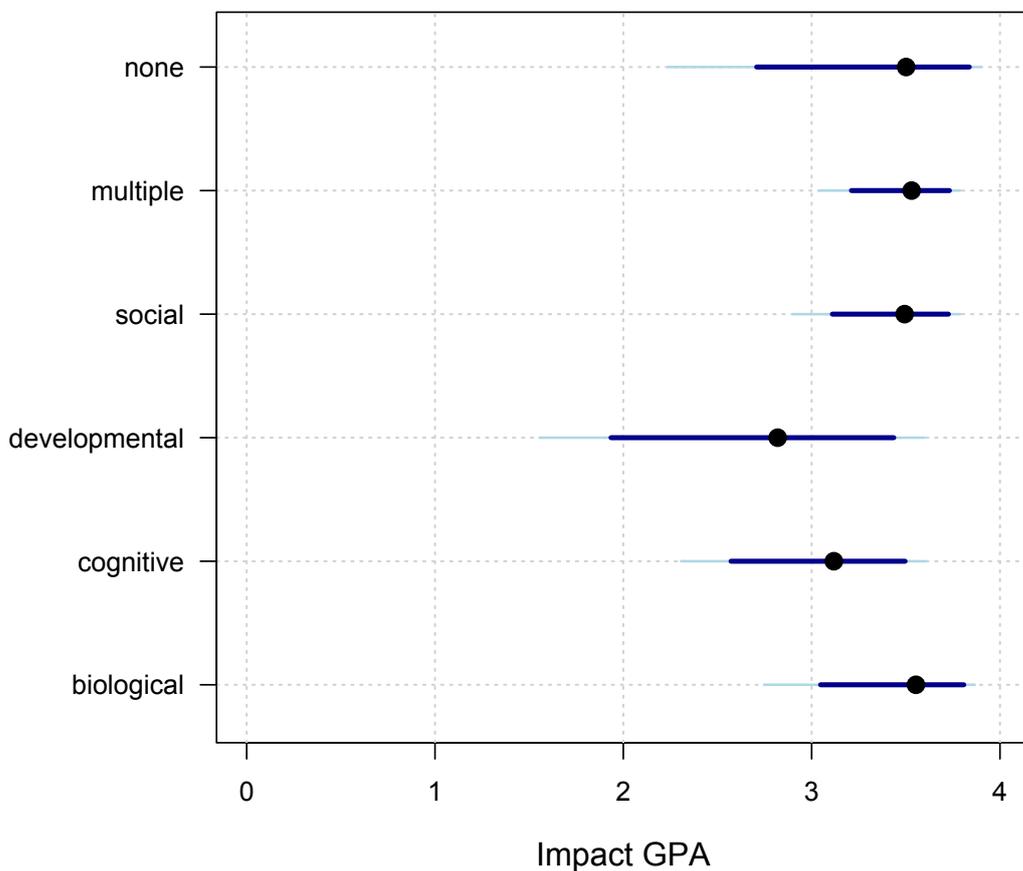


Figure 8: ...adjusting for output GPA, environment GPA and log(FTE)



less certainty about these estimates. In contrast sport and exercise psychology impact case studies are predicted to have low impact.

One concern in interpreting this pattern is that impact case study area may be confounded with other correlates of impact GPA. For example large submissions are qualitatively different from small submissions. Large submissions typically being two or more departments combined and often being linked to a large medical school.

Figure 6 shows how the pattern of predicted impact changes when output GPA, environment GPA and size of submission (log FTE) are accounted for. In this analysis the pattern is broadly similar, though now sport and exercise joins the group with middling, uncertain estimates and neuroscience lags behind the other areas. This suggests that the 'underperformance' of sport and exercise in the previous analysis is at least partly a consequence of sport and exercise case studies (which are relatively rare in the data set) being associated with weaker, smaller submissions.

A major caveat in interpreting these analyses is that few case studies fall in only a single core or applied area. Estimating the contribution of each area therefore runs into problems with multicollinearity – making it hard to tease apart the effect of each area (reducing statistical power) and perhaps making the interpretation of the effects difficult. For example, what does it mean to estimate the impact GPA for a neuroscience case study that isn't coded as biological, clinical or health? This issue is addressed in subsequent analyses looking separately at core and applied areas.

2.2.3 Analysis of core psychology areas

To avoid problems with collinearity this analysis focuses on the core areas of biological, cognitive, developmental and social psychology. Each case study was coded (prior to calculating averages at the submission level) as falling into to just one of these areas, into multiple areas or none (the latter possible, for example, for some medical research).

Figure 7 shows the predicted impact for a case study in each core area. This broadly matches the results from the earlier analysis – with the further observation that case studies drawing on two or more core areas a predicted to have high impact. The developmental psychology estimate is rather imprecise – largely because it is relatively infrequent as the sole core area (e.g. cognitive and developmental is a common combination). Figure 8 includes three covariates known to predict impact GPA and reveals a less differentiated pattern. These are the estimates for a typical unit (with average outputs and median FTE).

Broadly speaking all of the areas are similarly predictive of high impact after adjusting for strength and size of submission. Notably there is no sign that social psychology 'underperforms'.

2.2.4 Analysis of applied psychology areas

Figure 9 shows the predicted impact for applied areas (again separating out case studies with multiple or no applied area coded). Drawing on multiple applied areas, health psychology or forensic psychology predicts high impact scores. Estimates for clinical, occupational and neuropsychology are less certain, either because the case studies are rare (occupational) or rare as the only applied area. Educational psychology case studies are predicted to have slightly lower impact scores, but sport and exercise psychology case studies are predicted to have very low impact GPA.

Figure 9: Predicting impact GPA by applied area

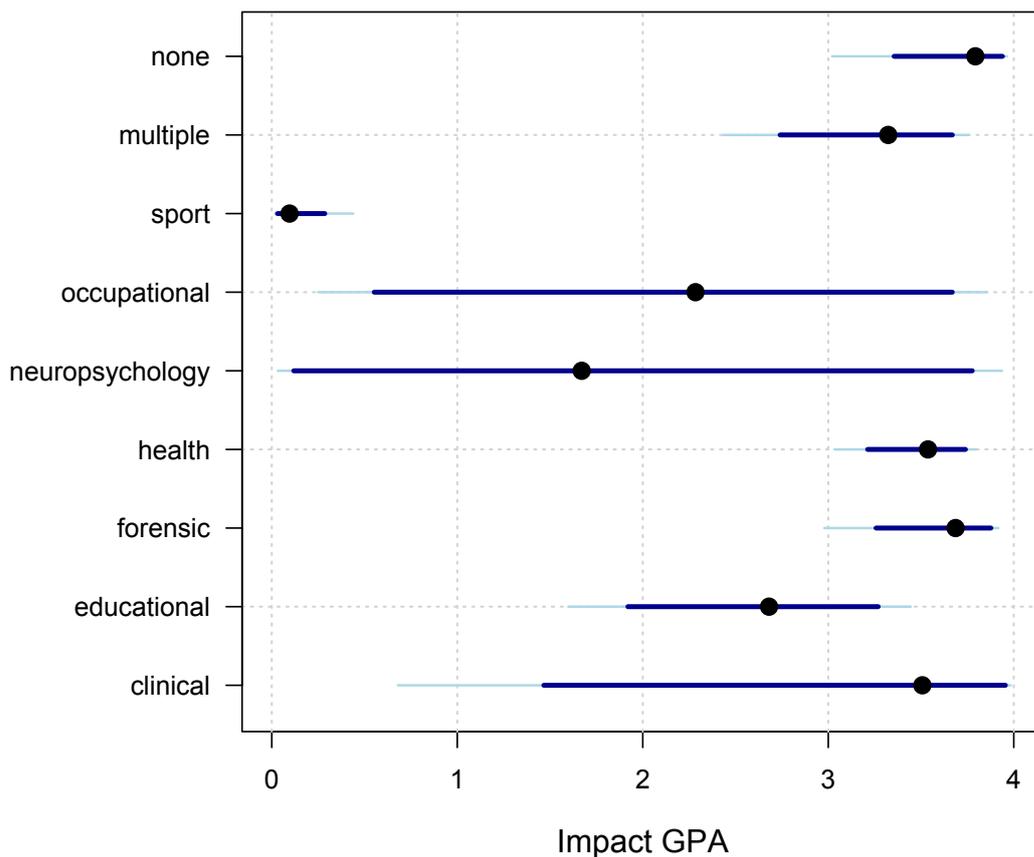
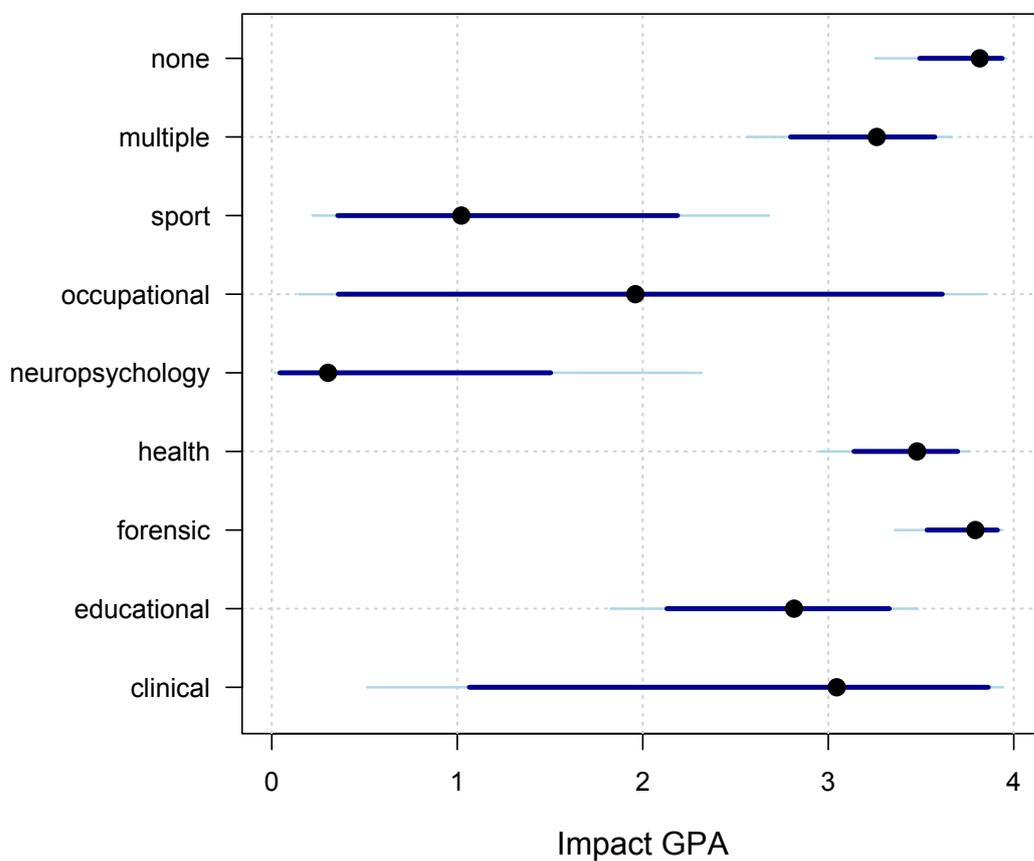


Figure 10: ...adjusting for output GPA, environment GPA and log(FTE)



This picture changes somewhat when accounting for strength and size of submission. Figure 10 shows a similar pattern but with higher (but still low in GPA terms) predicted impact for sport and exercise psychology case studies and lower impact GPA for neuropsychology. Again, these are the estimates for a typical unit (with average outputs and median FTE). Even taking into account that sport and exercise case studies are coming from weaker units there is some indication that they are expected to attract lower impact scores. For neuropsychology case studies it appears that they 'underperform' for a difference reason: they tend to come from stronger submissions and so the predicted impact GPA is below what one would expect of a strong submission.

2.2.5 Conclusions

The REF 2014 case studies offer a number of potential insights:

- It is clear that they draw on a very broad range of core areas of psychology with cognitive and social psychology each associated with nearly 40 per cent of case studies. In terms of applied area health is particularly prominent, but it may be hard to disentangle health psychology from interdisciplinary health research. Some areas, however, seem underrepresented – notably educational, sport and occupational psychology. This may reflect work by psychologists being returned to other UoAs.
- Many case studies draw on multiple core or applied areas and submissions with case studies that combine on multiple areas have high predicted impact GPA.
- The data suggest that all the core areas are more-or-less equally predictive of high impact GPA. In addition, areas such social psychology that are thought to be underrepresented in outputs are associated with strong impact case studies in the analyses reported here (and, as noted above, contribute to a high proportion of case studies).
- Applied areas show a more nuanced picture. There is evidence of strong performance (in terms of predicted impact GPA) for several areas – notably forensic psychology and health. For other areas there evidence across the analyses that sport, neuropsychology and neuroscience case studies predict lower impact GPA. These results need to be interpreted cautiously. There are very few sport psychology case studies and it is likely that they are atypical of sport psychology (e.g. because the strongest case studies are return as part of a sport science submission). The neuroscience and neuropsychology case predictions may also be atypical as much of the work in these areas would fall under multiple core or applied areas (e.g. cognitive neuroscience or clinical neuropsychology).

The British Psychological Society

St Andrews House, 48 Princess Road East, Leicester LE1 7DR, UK

Tel: 0116 254 9568 Fax 0116 227 1314 Email: mail@bps.org.uk Website: www.bps.org.uk