Quality & Quantity (2024) 58:829–848
https://doi.org/10.1007/s11135-023-01669-w

Charting fields and spaces quantitatively: from multiple correspondence analysis to categorical principal components analysis

Will Atkinson[1]

Accepted: 12 April 2023 / Published online: 28 April 2023

Abstract
Multiple correspondence analysis (MCA) has started to gain popularity within sociology as a method of mapping 'fields' and 'social spaces' in the style of Pierre Bourdieu, its capacity to document multidimensional geometric relationships within data being a snug fit for the relational mode of thought he championed. There is a risk, however, of over-relying on MCA when the data suggest alternative methods and, as a result, drawing unsound conclusions. As a case in point, I take a recent analysis of political attitudes in the UK using MCA that drew bold inferences about the relationship with social class and reanalyse the same data with categorical principal components analysis (CatPCA). The results suggest the opposite conclusion to what was originally argued. I thus urge greater methodological flexibility and openness among those wishing to chart fields and social spaces and, more specifically, I make a case for CatPCA as a tool of geometric data analysis.

Keywords Fields · Multiple correspondence analysis · Categorical principal components analysis · Geometric data analysis · Horseshoe effect

1 Introduction

Pierre Bourdieu pioneered the use of multiple correspondence analysis (MCA) within sociology as a technique for documenting the structure of 'fields' and 'social spaces' with quantitative data. Embedded within a French tradition of 'geometric data analysis' (GDA) founded by Jean-Paul Benzécri and harmonising with Bourdieu's relational mode of thought, but chafing with mainstream preferences for regression analysis (Bry et al. 2015), the technique rested on the margins of the discipline for some time. Over the last 15 years or so, however, there has been growing interest in using MCA and GDA to follow up on Bourdieu's project

✉ Will Atkinson
w.atkinson@bristol.ac.uk

[1] School of Sociology, Politics and International Studies, University of Bristol, 11 Priory Road, Bristol BS8 1TU, UK

and examine the structure of fields and social spaces across various nations in the 21st Century. This has yielded a productive and revealing burst of activity in studies of class, culture and politics in particular.

There is a risk, however, that some scholars inspired by Bourdieu and seeking to document fields and social spaces might over-rely on MCA when other techniques could be more suitable for the data at hand. MCA is, after all, best suited to multiple nominal variables – categorical variables with no obvious rank order – and when other types of variables are analysed it may be preferable to use alternative methods that nonetheless produce multidimensional spaces amenable to geometric data analysis. More than that: using MCA and MCA alone may generate misleading results and erroneous conclusions. This is the case I will make in this paper, taking as my example a recent study of the space of political attitudes and its relationship with social class (Ibrahim and Lindell, 2020). I reanalyse the same data using categorical principal components analysis (CatPCA) and produce results, and conclusions, quite at odds with those of the original authors. In short, while the original authors say the importance of class is limited and unidimensional, I demonstrate it to be substantial and multidimensional. I thus urge greater methodological flexibility among those interested in documenting fields and social spaces and make a case for CatPCA as a tool of GDA.

2 Trajectories of MCA

MCA as a technique is designed to uncover underlying relationships within quantitative data. More specifically, it looks for associations between categories of multiple variables, detecting, for example, that people who say X are also likely to say Y, but also that some people who say X and Y are likely to say A too while others say B instead. It does this by breaking down the total variance or inertia within the data into a series of axes – sometimes described as dimensions or factors – accounting for descending proportions (denoted by eigenvalues), the patterns of association and opposition between variable categories thus being translated into distances within a space. X and Y will be close together on one axis, but opposed to Z, while X, Y and Z might be close together on a second axis but A and B polarised. The analyst retains for interpretation only those first n axes that together account for a satisfactory proportion of total inertia, also paying attention to any substantial drop-off in inertia rates between axes.

MCA is similar to linear principal components analysis (PCA), the only differences being that (i) MCA works with categorical variables while PCA works with continuous variables and (ii) MCA arrays categories of variables separately within the space while PCA works with whole variables via correlation or covariance. MCA, focussing on category centroid coordinates, thus allows for non-linear relationships to emerge, as ‘middle’ categories might display very particular associations, whereas PCA focuses on *vectors*, or the general directions and distances of travel of variables as wholes within the space.

MCA has been developed independently by different statistical traditions (DiFranco, 2016). The Leiden school of psychometry elaborated it under the label of ‘homogeneity analysis’ in the 1970s and 1980s, alongside PCA and other techniques, as a method of data reduction that can produce axes definable as scales for subsequent regression analysis (Gifi 1990). In France, MCA was spearheaded by Benzécri and evolved, in the work of Ludovic

Lebart, Brigitte Le Roux and Henri Rouanet, into the broader programme of GDA, the core interest of which is the geometric properties of the space itself, that is, the distances and directions of deviation between variables/categories or cases (see esp. Lebart et al. 1984, 2006; Le Roux and Rouanet 2004, 2010; Le Roux 2014; Le Roux et al. 2020). An extra layer of analysis was enabled by the distinction between 'active' variables – variables used to construct the model – and 'supplementary' variables – variables not used to construct the space. This is because positions and dispersion along axes of supplementary categories can reveal correspondences and homologies between 'outcomes' and possible causal factors. Those advancing the GDA tradition have, moreover, established a number of statistical tests and guides for interpreting these correspondences. Lebart et al. (2006), for example, devised a test similar to the one-sample t-test to determine whether or not a category's location on an axis is important (the congenial language of 'atypicality' is used for a similar test developed by Le Roux and Rouanet 2004), and a distance between factor coordinates is deemed notable when it is 0.4 or more units or standardised deviations (SDs) on an axis and large if it 1.0 or more (Le Roux et al. 2020). Univariate analysis of variance with an axis as the dependant variable, and the related eta-squared figure denoting explained variance on the axis (on a 0–1 scale), also allows one to adjudicate the extent to which a variable is associated with an axis given inter- and intra-category dispersion (Hjellbrekke 2019).[1]

3 Bourdieu and MCA

Bourdieu came into contact with Benzécri and Lebart in the 1970s and, in later years, worked directly with Le Roux and Rouanet (Bourdieu 2008). He believed MCA to be a good fit for his relational mode of thought centred around a conception of social structures in terms of 'fields' and 'social spaces' (Bourdieu 1984; Bourdieu and Wacquant 1992; Bourdieu et al. 1991). A field is an arena of social struggle between a set of agents over certain forms of power or 'capital', possession of different types and amounts of which constitute the structure of the field and the conditions of possibility for specific perceptions and practices. It is noticeable, in fact, that Bourdieu's theorisation of fields as relational structures took on a specific form after his contact with Benzécri's school in the 1970s: a more spatialised or geometric form, which harmonised with a structuralist vision organised around oppositions, directions and distances rather than interactions and which even tended to borrow operations and terms from the statistical methods such as 'inertia', 'weight' and so on. Conceptual development and methodological development went hand in hand, in other words, just as Gaston Bachelard, Bourdieu's prime epistemological influence, would have it.

Examples of fields documented by Bourdieu and his colleagues using MCA include the academic field (Bourdieu 1988), the field of CEOs (Bourdieu 1996) and the field of publishing houses (Bourdieu 2008). The overall class structure of a society is also a field, though Bourdieu (1984) preferred to call this the 'social space', and is structured by possession of economic capital (money, assets), cultural capital (credentials, parental education) and

[1] Meanwhile Michael Greenacre (2017), a former student of Benzecri, has proposed several technical adjustments to MCA, including 'joint correspondence analysis' and axis rescaling, but these have not (yet) been accepted by most practitioners of French GDA. Together with Jörg Blasius, moreover, Greenacre has led efforts to bridge statistical traditions (see e.g. Greenacre and Blasius 2006; Blasius and Greenacre 2014), but, in practical terms, the researchers carry on with their own assumptions, tests and even notations.

social capital (connections, memberships). The primary axis of the social space was held to be total capital *volume*, while a secondary axis distinguished people on account of their capital *composition*, i.e. whether it is predominantly economic or cultural in constitution. This represents a break with unidimensional visions of stratification, whether measured via socio-economic scales or occupational schemas. Bourdieu did not map the French social space using MCA, however. Instead he mapped out the space of lifestyles: the space of goods, practices and tastes acting as symbols of class position. More accurately he mapped out two, one for the dominant class and one for the intermediate class. He could then explore the correspondences with social positions, or the homology with the social space, by projecting indicators of capitals, as supplementary variables, into the models of the lifestyle spaces, though he was working before the aforementioned tests and benchmarks were established.

4 Subsequent uses and limits of MCA

After a period of relative neglect, MCA has, over the last 15 years, started to penetrate mainstream sociology. Scholars interested in testing or updating Bourdieu's ideas, mostly but not exclusively across Europe, have started to adopt it (see Robson and Sanders 2009; Grenfell and Lebaron 2013; Blasius et al. 2020). For the most part this has been mediated by the continued influence of Le Roux and Rouanet, meaning the majority – there are some exceptions – implement the specifically French GDA-style approach to MCA. Especially prominent have been studies of national social spaces (e.g. Prieur et al. 2008, Rosenlund 2009, Flemmen et al. 2018; Atkinson 2022a) and lifestyle spaces (Bennett et al., 2008; Le Roux et al. 2008; Blasius and Muchlichen, 2010; Flemmen et al. 2019; Coulangeon and Duval 2015; Schmitz 2017; Hanquinet and Savage 2018; Atkinson 2017, 2021, 2022b; Atkinson and Marzec 2023), but another strand of work has also examined the space of political attitudes and its relationship with class (Harrits et al. 2010; Flemmen 2014; Flemmen and Haakestad, 2018; Barth and Schmitz 2018; Jarness et al. 2019; Atkinson, 2017). The logic in the last case, building on arguments and sketchy analyses by Bourdieu (1984) himself, is that political attitudes form a relational system symbolising class ethos in a similar manner to lifestyles. MCA offers an alternative to the common use of factor analysis in the study of political attitudes, as in the work of Ronald Inglehart (1977, 1990, 1997; Norris and Inglehart 2019) for example, and the usual finding is that political attitudes fall along two axes: one distinguishing 'old' left/right-wing views on economic matters (e.g. income redistribution, collective ownership, industrial relations) and another distinguishing 'new' liberal/traditional-authoritarian views on 'moral' or non-material matters (e.g. homosexuality, immigration, defence). Summarising crudely, high/low cultural capital corresponds with the liberalism/authoritarian axis and high/low economic capital corresponds with the right/left-wing economics dimension, but this means that capital volume and capital composition are associated with different shades, combinations and polarisations of views.

At stake in many of these Bourdieu-inspired studies is not just the relative importance of social class but the status of capital composition as a principle of social, cultural and political differentiation. This is important from a theoretical point of view because it relates to whether and to what extent class structures are multidimensional – and thus how well they are approximated by catch-all hierarchal measures like occupational schemes – and if Bourdieu's thesis might overplay the cultural/economic polarisation because of a Francocentric overestimation of the importance of intellectuals (e.g. Lamont 1992; Savage 2010, 2021). Some analyses using MCA have, indeed, questioned it. Several studies of lifestyles, for example, have suggested capital composition is subordinate to age or gender in polarising tastes and practice, if it appears at all (e.g. Bennett et al., 2008; Le Roux et al. 2008; Börjesson, 2015). In these cases, the conclusions rest on the specific selection of variables for analysis, and when alternative variables are used – arguably more suited to testing the capital composition thesis – Bourdieu's model is more or less confirmed and updated, including in the very same national cases where it is apparently absent (see e.g. Flemmen et al. 2019; Atkinson 2017, 2021).

In other cases, however, the capital composition principle is questioned because the MCA model suggests a unidimensional solution. More precisely, the categories and cases on the plane of the first two axes are arranged in a parabolic U-curve known as the horseshoe or Guttman effect, indicating strong intercorrelation between the active variables. In practical terms, the first axis usually opposes high/low-style responses to questions while the second axis opposes extreme/middle responses. If the space in question is a model of the social space – as in Lemel and Katz-Gerro (2015), Doolan and Tonkovic (2021) or Marzec (2019), for example – then the first axis distinguishes those with high/low capital while the second axis distinguishes those with high *and* low capital from those with middling levels of capital.

The Guttman effect is not necessarily a statistical artefact, but neither is it especially interesting. The usual response is to seek to go beyond it and find other dimensions of greater sociological interest (Hjellbrekke 2019: 96). This can be done by focussing on a third axis (e.g. Harrits et al. 2010; Marzec 2019), aggregating variable categories (e.g. Meuleman and Savage 2013) or employing forms of fuzzy coding such as 'doubling' (Greenacre 2017). Sometimes, however, the effect can account for so much inertia in the model as to make interpreting a third axis hard to justify from a statistical point of view, or it can persist even after extensive recoding, which can itself become distortive or reduce differentiating power in the model (or produce artificial models in the case of doubling). The effect can be especially intractable if the variables analysed are Likert scale items. After all, the Guttman effect efficiently reveals that the variables selected for analysis follow an *ordinal* structure – there is a known low-to-high rank order to responses – rather than a strictly multiple nominal structure (Le Roux and Rouanet 2004: 220-1). In this case some researchers would turn to a different method designed to handle ordinal variables: CatPCA.

CatPCA, also known as nonlinear PCA (NLPCA), was developed by the Leiden School and, in technical terms, operates in precisely the same way as MCA and linear PCA (Di Franco 2016): it constructs a space defined by axes of varying contribution to total model inertia that can then be subjected to follow-on tests of supplementary variables in the style of GDA. The difference is that CatPCA can handle any and all types of data simultaneously – continuous and categorical data – and, crucially, one can specify beforehand if categorical variables follow an ordinal or simple/multiple nominal structure. If a variable is set to follow an ordinal structure, then the low-to-high structure is assumed and the variable is constrained to follow a vector through the centre of the space, thus erasing the possibility of it forming a horseshoe.[2] This is accomplished via optimal scaling. Categorises are quantified, i.e. assigned numerical values on a scale, in such a way as to maximise explained inertia in

[2] One can also insert a specified number of 'splines' into an ordinal variable, which allow the vector to curve, but this is most suited to variables with many categories.

the model. One can then inspect transformation plots for each variable, plotting category quantifications in a line graph, to adjudicate whether a variable is indeed best treated as nominal or ordinal, to see where the key points of differentiation within the variable are and to check model stability (see Atkinson 2020). CatPCA as a method is therefore much more flexible than either MCA or linear PCA and allows an efficient escape from the Guttman effect. The downsides are that it is not as simple since the optimisation phase requires the analyst to iteratively inspect, determine and justify the best measurement level of each variable, and that resultant models are sensitive to such decisions. There is also the risk that a horseshoe effect is erased without acknowledgment of its existence or strength if variables are assumed to be, and set as, ordinal from the outset. That is as undesirable as being unable to move past the effect. Starting from the multiple nominal level of categorical measurement, perhaps with a straightforward MCA, is thus prudent before proceeding to explore other options.

There are a few minor operational differences between MCA and CatPCA. One is that axis retention in CatPCA is based on slightly different criteria and takes place only after examination of solutions of varying pre-set dimensionality. These criteria include the Kaiser principle (eigenvalue > 1), very good total variance accounted for on each variable (> 0.4 according to Comrey 1973) and satisfactory explained inertia in the model (e.g. >50%), as well as the usual considerations of sociological interpretability and parsimony. Models can be rotated, in the style of linear PCA, to modify differences and axis inertia, but many working in the French GDA tradition criticise rotation on the grounds that it introduces artificial distortion and produces a less proximate model of reality (Rosenlund 2009: 154) and it is not advocated. After all, the point of rotation is often to maximise axis orthogonality for subsequent regression analysis while the focus of GDA is on the base dimensions, distances and directions of deviance. For this reason, Cronbach's alpha is of little interest. A risk of the French approach, though, is that one can sometimes be left with axes that represent skewed proxies for hypothesised or theoretically meaningful axes, which then effects the capacity to perform subsequent tests satisfactorily.[3]

Another difference between the methods is that, in judging the importance of a variable to an axis in CatPCA, one pays attention to the factor loading, or the correlation of the variable with the axis, with anything greater than +/-0.4 being considered important. However, one can calculate from this the relative contribution of the variable to the axis variance and highlight above-average ('explicative') contributions, matching the common logic of MCA.[4] Finally, centroid coordinates of variable categories in CatPCA are mean scores on axes rather than factor coordinates. Lebart et al's (2006) test values for notability can still be calculated, but instead of SDs one can use independent (Student's) t-tests to determine whether the distances between points are important or not, so long as this is understood as a benchmark for assessing the notability of distances rather than a tool of hypothesis testing.[5] This can be paired with Cohen's d to estimate the 'effect size' in a manner comparable to SDs: anything between 0.2 and 0.5 is considered small (but not negligible), anything between 0.5 and 0.8 is typically considered a 'medium' effect and anything over 0.8 is

[3] A compromise solution was proposed to this problem in Atkinson (2020), though how satisfactory it was is open to debate.

[4] The relative contribution to inertia is equal to the squared factor loading divided by the axis eigenvalue.

[5] This difference in centroid coordinates was not sufficiently appreciated in Atkinson (2020), where the importance of distances in the CatPCA spaces presented there are therefore underestimated.

considered 'large', though of course these are points on a scale rather than fixed categorical thresholds (Cohen 1988: s.2.2.3).

CatPCA is employed readily within political science and sometimes within sociology (e.g. Maire 2021). It is not, however, currently considered part of the French GDA toolkit, even though linear PCA is. It is absent, for example, from the major texts of the tradition, including Le Roux and Rouanet (2004), Lebart et al. (2006) and, more recently, Husson et al. (2017). The disconnect is even perpetuated by software differences: SPAD, a French software package for conducting MCA and other GDA methods recommended by Le Roux and Rouanet (2010) and Hjellbrekke (2019), does not currently contain a module on CatPCA. CatPCA, on the other hand, is available in, among other software packages, R, Stata and SPSS, the last of these containing an MCA module designed by the Leiden School but explicitly warned against by Le Roux and Rouanet (2010).[6]

As a consequence, CatPCA is not typically deployed by the Bourdieu-inspired wishing to construct and explore statistical models of fields or spaces. An exception is Joye et al. (2020), whose mission to broaden the methodological mindset of Bourdieusians and encourage experiments and comparisons with different spatialising techniques is similar to that expressed here (see also Rosenlund 2009), though their focus is only on concordance of solutions and they do not explore geometric properties of the spaces. Most, however, stick resolutely to MCA, and where the horseshoe effect occurs it is sometimes – as in Lemel and Katz-Gerro (2015) or Doolan and Tonkovic (2021) – simply reported as an approximation of social reality. That is not inadmissible – the effect does demonstrate the importance of the middle/extremes opposition – but it is problematic to then leap from the model to reality and conclude that any other dimensions of difference are irrelevant or sociologically marginal. It may simply be that the MCA model is telling us that all/most of the variables follow an ordinal structure and should be treated as such. After all, MCA was designed for multiple nominal variables and is ideal for variables such as favourite restaurant, industry of occupation or political party of choice. These can even be mixed with ordinal variables without much disturbance – a Guttman effect may appear on a lower-order axis, for example (as, indeed, it did for Bourdieu, [1984], Bennett et al. [2008] and Atkinson [2022a]), if at all – but the more ordinal variables there are, to the limit point of having exclusively ordinal variables, the more the chances of obtaining a Guttman effect in the principal plane increase. Perhaps for this reason, Rosenlund (2015) expressly recommended including a multiple nominal measure of occupation or industry when constructing models of social spaces with MCA, since many measures of capital like education level or income do often follow ordinal structures, though if there are no such variables available in the dataset at hand then that advice founders. So why not turn to CatCPA instead? Possibly because MCA is considered exhaustive, in the minds of many of those inspired by Bourdieu, of statistical methods of constructing and exploring spaces, whether because of lack of familiarity with other methods or aversion to techniques associated with very different traditions. Yet there may be more than one way to skin the proverbial cat. There may even be *better* ways to flay the feline.

[6] In what follows, I will use SPAD for MCA and SPSS for CatPCA, though graphical representations for the latter have been generated in Excel because of low functionality in SPSS. Test values for the CatPCA are generated by importing saved axes into SPAD and running biplots.

5 The British space of political attitudes

Elsewhere I have suggested the utility of using CatPCA for constructing models of national social spaces when MCA struggles to get beyond horseshoes (Atkinson 2020, 2022c), but here I want to demonstrate its usefulness in relation to studying the space of political attitudes. As a case study I will take the recent analysis of British attitudes produced by Lindell and Ibrahim (2020). This case has been selected because of Lindell and Ibrahim's commitment to Bourdieu-style MCA, the boldness of their conclusions, the stark contrast with previous findings, including in the same national case (Atkinson 2017), the public availability of their chosen data source and the welcome openness of the authors to further studies on the dimensionality of political attitudes. It is not a comment on the facility of the authors with either MCA or British politics or on the quality of their work.

Lindell and Ibrahim use the following five-point Likert scale variables denoting left/right politics and liberal/authoritarian attitudes from the 2016 sweep of the European Social Survey (ESS) to construct an MCA model of the post-Brexit space of political attitudes and examine the relationship with class and other factors:

- Self-placement on left-right scale.
- Government should reduce differences in income levels (Gov inc).
- There should be benefits for parents to combine work and family even if it means higher taxes (Benefits).
- Against or in favour of a basic income scheme (Bas inc).
- Homosexual couples should have the right to adopt children (Adopt).
- European Union: European unification should go further or has gone too far (EU).
- Allow many/few immigrants of different race/ethnic group from the majority into the country (Allow).
- Increase taxes on fossil fuels to reduce climate change (Tax foss).

Their model suggested that left/right and liberal/authoritarian attitudes are strongly correlated along the first axis and cross-cut on the second axis only by intensity of opinion – in other words, they uncover a Guttman effect (cf. Majima and Savage, 2008). The authors stated that, when faced with this result, they tried recoding the five-point scales into three-point scales but produced the same outcome, opting to retain the five-point scales apparently because of greater proximity to reality. They declined to discuss the third axis, even though it had an inertia rate comparable to the second dimension and substantially higher than the fourth dimension (representing drop-off), because it is described as a 'variant' of the first axis. Ultimately, they concluded, intensity of opinion is a real and powerful feature of British political attitudes crowding out any other secondary polarisations. Scrutinizing supplementary variables, they conclude with surprise – since it contradicts established patterns – that class is relatively unimportant for shaping political attitudes and, on top of that, any effects of capital composition are entirely absent. They suggest this may reflect the social realities of contemporary Britain as opposed to Bourdieu's France.

One can raise questions about whether the active variables are really the best to tap into the different dimensions of political attitudes, and the use of five-point scales with low-frequency extremity categories risks model instability, i.e. attributing small-n categories disproportionate power in determining axes (Hjellbrekke 2019). We will put those issues to one side, however. More problematic is the reliance on basic unidimensional indicators of class position, especially the UK National-Statistics Socio-Economic Classification of occupations (NS-SeC). This is an insufficient proxy for any posited multidimensionality of the social space and an alternative classification of occupations designed to approximate capital volume and capital composition, and validated via MCA, will be used instead (Fig. 1) (Atkinson 2017, 2022c; Atkinson and Rosenlund 2014).

I have taken the liberty of obtaining the 2016 ESS dataset and replicating Lindell and Ibrahim's analysis as closely as possible. The results obtained from the replication were not exactly the same as theirs, which might have something to do with the unspecified 'missing values' in Lindell and Ibrahim's analysis (missing values were treated as passive in my analysis rather than excluded). Still, relative frequencies and the general distribution of categories in the plane of Axes 1 and 2 are almost identical (Fig. 2). Inclusion of the Bourdieusian class scheme already reveals a stronger relationship with political attitudes than originally supposed: coordinates of different class fractions are more dispersed, important according to tests and notably distanced from others in terms of SDs (Fig. 3, Table 1). We learn that the cultural dominant class fraction is most left-liberal and that technicians, skilled workers and lower managers/proprietors (LMPs) are most right-authoritarian, but also that caring/sales workers and cultural intermediaries are associated with less 'intense' opinions.

The third axis is not, in fact, merely a variant of the first. It operates to split the liberal-left pole of Axis 1 into a liberal/pro-EU quadrant (bottom right) and an economically egalitarian quadrant (top right) and to split the authoritarian-right pole into a mild version (bottom left, at the same pole as liberalism on Axis 3) and a more extreme version (top left, at the same pole as economic egalitarianism) (Fig. 4). This is not all that far from what has been found in other studies. However, the modified inertia rates in the present model, using Benzécri's (1992) formula to correct for underestimation, are different from Lindell and Ibrahim's: in this version, the first axis bears a rate of 68% and the second a rate of 23%, followed by

Capital volume +

Cultural capital + | | *Economic capital +*

Cultural dominant (e.g. cultural producers, intellectuals, teachers)	**Professions** (e.g. doctors, lawyers, architects) **White-collar workers** (e.g. software developers, surveyors, HR officers)	**Business executives** (e.g. managers and CEOs in finance and manufacturing)
Cultural intermediaries (e.g. nurses, therapists, youth workers)	**Technicians** (e.g. engineers, lab technicians) **Administrators** (e.g. secretaries, bookkeepers, clerks)	**Lower managers/proprietors (LMPs)** (e.g. garage managers, hotel managers, shopkeepers)
Caring services (e.g. dental nurses, hairdressers, travel agents)	**Sales workers** (e.g. retail assistants, cashiers, window dressers) **Manual workers** (e.g. cleaners, shelf-stackers, assembly line workers)	**Skilled trades** (e.g. mechanics, electricians, plumbers)

Capital volume -

Fig. 1 A class scheme to approximate the British social space

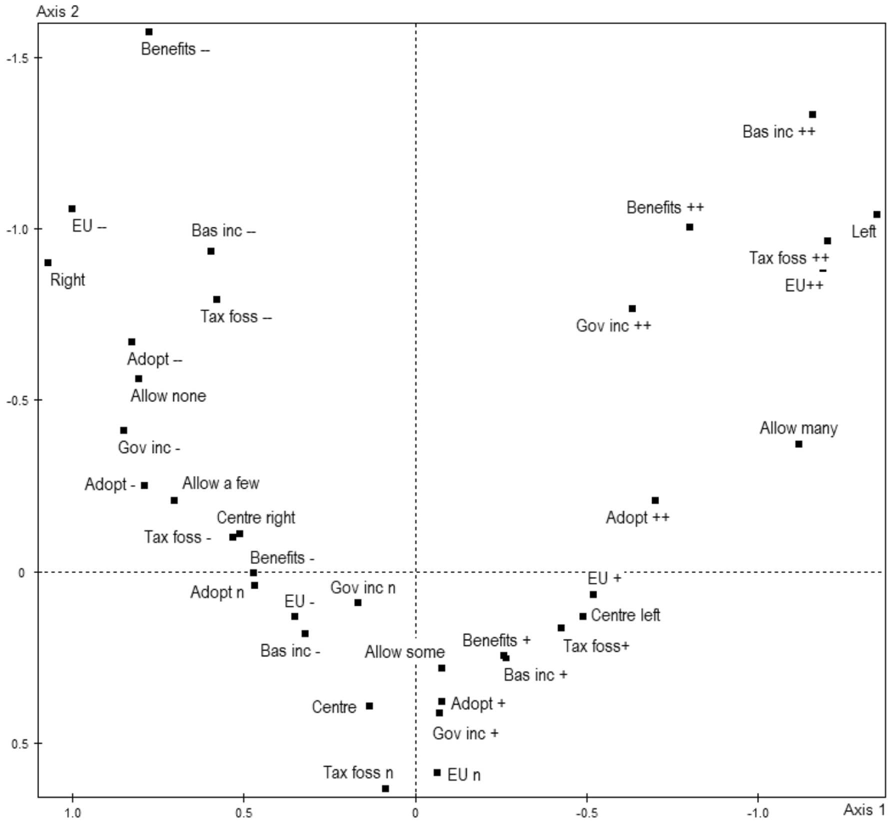

Fig. 2 MCA space of political attitudes, axes 1 and 2

just 4% on the third. In this instance, then, there is certainly a good argument for forgetting about the third axis. Many of the small-n extremity categories also have relatively powerful contributions on one or more axis, attributing a lot of structuring power to relatively few people and confirming earlier reservations about stability. Aggregating all the variables into three categories remedies this but has the effect of producing a first axis with a modified inertia rate of 84%, making the model truly unidimensional (the second dimension still, as Lindell and Ibrahim found, produces a parabolic curve).[7]

[7] One could argue, with Greenacre (2017), that Benzecri's modification overestimates inertia, and that applying Greenacre's adjustment would produce more modest relative inertia rates on the first axes. However, most researchers working in the Bourdieusian/French tradition – including Lindell and Ibrahim – use only Benzecri's correction so I will follow suit here.

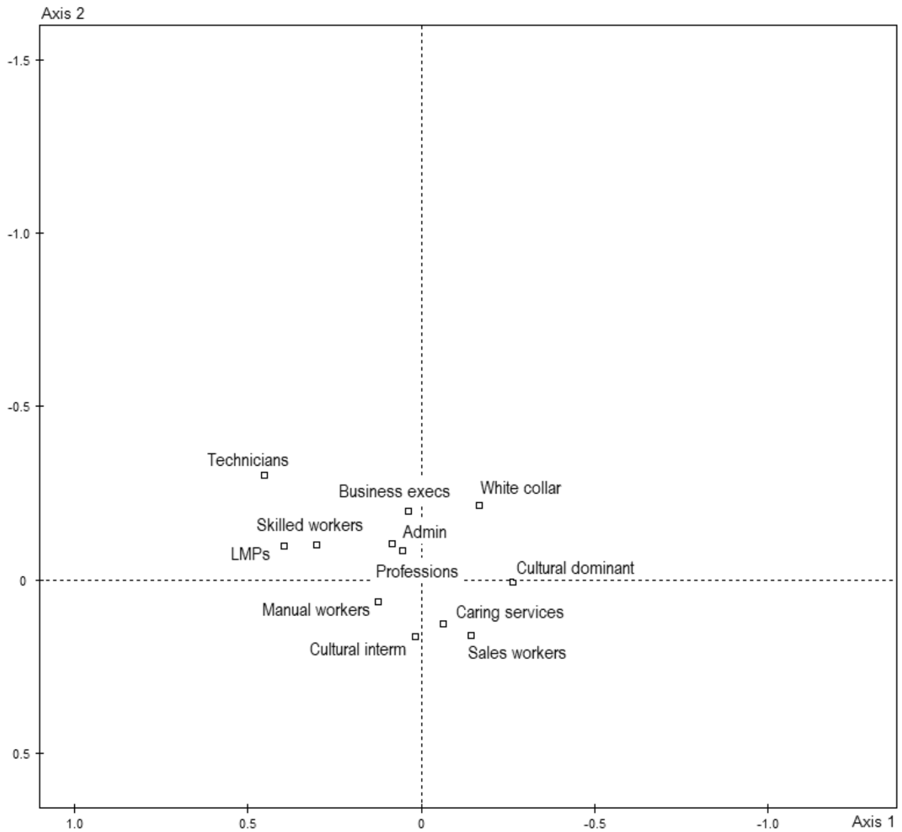

Fig. 3 Class categories in the space of political attitudes

Table 1 Class fractions in the MCA space

	%	Test value	
		Axis 1	Axis 2
Business execs	4.1	0.33	-1.74
Professions	3.4	-1.33	2.15*
White collar	9.1	-2.33*	-3.59**
Cultural dominant	9.3	-3.35**	0.12
LMPs	5.2	3.81**	-0.93
Technicians	5.1	1.33	-2.80**
Administrators	10.2	2.16*	-0.19
Cultural intermediate	5.2	-0.25	0.89
Skilled workers	10.2	3.82**	-1.10
Manual workers	16.8	2.44*	0.49
Sales workers	11.4	-2.14*	2.64**
Caring services	10.1	-0.92	1.78

*p < 0.05, **p < 0.01.

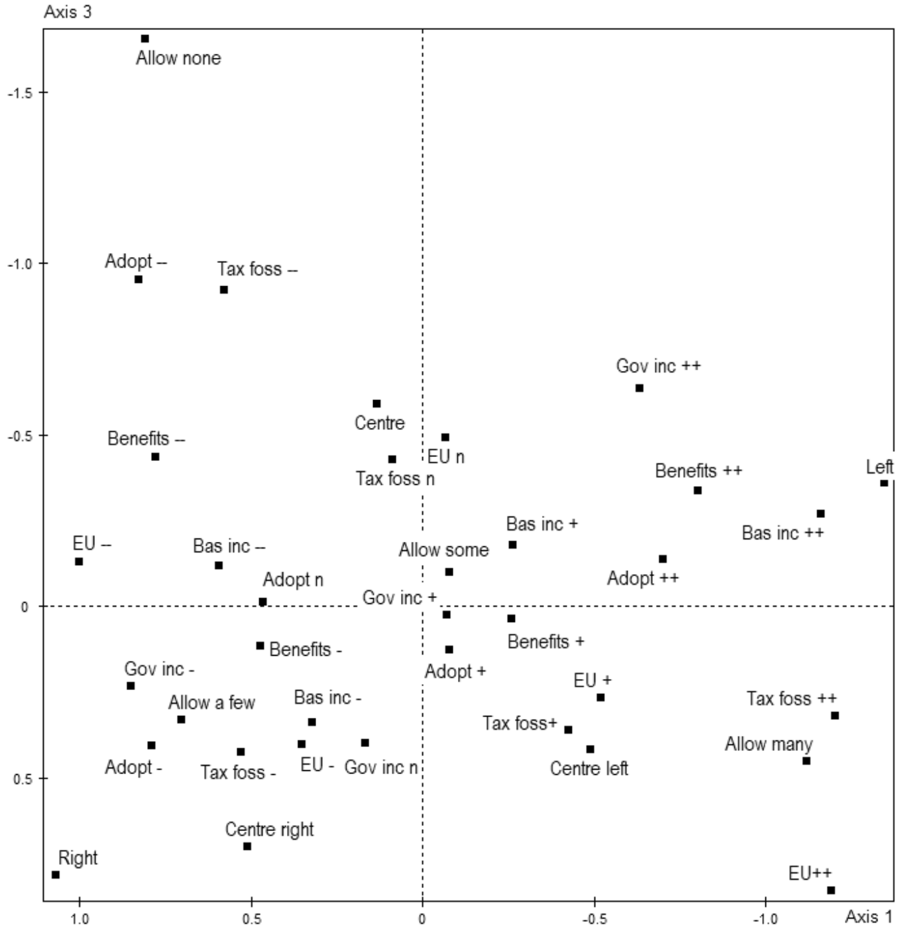

Fig. 4 MCA space of political attitudes, axes 1 and 3

6 Splitting the horseshoe

We can accept, then, that the MCA reveals that intensity of opinion matters in structuring political attitudes. To say that any other principles of difference are sociologically marginal or irrelevant, however, is too hasty and possibly to mistreat the data on the basis of what they are telling us in the MCA. Lindell and Ibrahim quote Benzécri to the effect that the model should fit the data, not the other way around, but if the data are clearly revealed to be ordinal, then should we not construct a model fitting that structure? We are not hiding or downplaying the middle/extreme opposition – far from it, since I have also argued for its importance on similar grounds (Atkinson 2017) – but it is worth exploring other methods specifically designed to look beyond the extreme/middle divide, i.e. CatPCA. MCA is then a necessary preliminary to CatPCA: it establishes the ordinal nature of responses, and that intensity of opinion (or some other middle/extreme divide) matters, and if third axes or moderate recoding are off the table then we move on to other methods to explore salient polarisations obscured by the first method.

Running a CatPCA on exactly the same data, with the variables set as ordinal, certainly provides a very different view. Three axes are extracted on the basis of the criteria set out earlier (Table 2).[8] The first of these, judging from factor loadings and relative contributions of variables, is the same opposition of right/authoritarian and left/liberal that emerged in the MCA, confirming that this is indeed the key dimension of difference structuring British politics. We do see, however, that attitudes toward classic materialist issues of redistribution and welfare are less important to the axis than views on immigration, Europe, sexuality and fossil fuels, suggesting the liberal/anti-liberal polarity is the primary structuring feature of political attitudes in the sample. The second axis bears an opposition between those who are right-wing on material matters but liberal on ‘new’ politics – especially immigration – and those who are left-wing on old politics but anti-migration. Looking at vector coordinates in the plane of axes one and two (Fig. 5), we see four distinct quadrants: left-wing (top left), right-wing (bottom right), liberal (bottom left) and anti-liberal (top right). This structure approximates what has been found in previous Bourdieu-inspired studies, including in the UK (Harrits et al. 2010; Flemmen and Haakestad, 2018; Jarness et al. 2019; Atkinson 2017).[9] The third axis, for its part, is structured around an opposition between those who are against government redistribution (and who self-identify as right-wing), but nonetheless in favour of benefits to get people working even if it means higher taxes, and those of the opposite view, who also oppose taxing fossil fuels.

Focusing on the plane of Axes 1 and 2, projection of indicators of capital and class position as supplementary variables reveals a very different picture from what emerged in

Table 2 Properties of the CatPCA space

	Axis 1	Axis 2	Axis 3
Eigenvalue	2.162	1.174	1.025
Proportion of inertia	27.0	14.7	12.8
Factor loadings (% inertia)			
Right wing	0.58 (15)	-0.20 (3)	0.45 (19)
EU +	-0.58 (16)	-0.27 (6)	0.13 (2)
Homosexual adoption -	0.59 (16)	0.21 (4)	0.21 (4)
Immigration -	0.62 (18)	0.42 (15)	-0.03 (0)
Tax on fossil fuels -	0.52 (13)	0.33 (9)	-0.26 (7)
Gov redist. -	0.43 (9)	-0.54 (25)	0.47 (22)
Benefits +	-0.36 (6)	0.45 (17)	0.60 (36)
Basic income +	-0.42 (8)	0.49 (21)	0.33 (10)

Notes: Loadings greater than +/-0.4 are considered important to an axis. Relative contributions to inertia greater than 12.5% are considered explicative. Directionality of the questions is indicated with – (against) or + (for).

[8] The model was bootstrapped to check stability. This operation generates 95% confidence intervals for all relevant output and thus allows one to judge the quality of axes.

[9] The vector coordinates also reveal the distal positioning of the low-frequency ‘strongly agree’ categories for basic income (rel. freq.=6%) and benefits (7%). Their strong effects are confirmed by examining category quantifications in transformation plots. Recoding just these two variables did not eliminate the horseshoe effect in the MCA space, however, nor did it effect the overall structure of the CatPCA space.

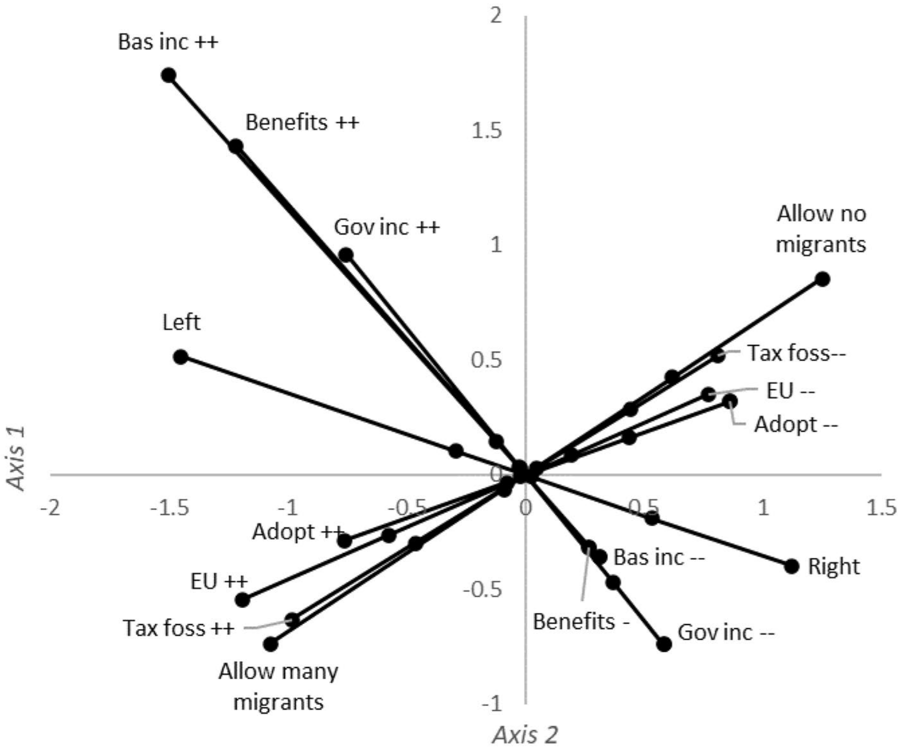

Fig. 5 Vector coordinates, axes 1 and 2

Lindell and Ibrahim's paper (Fig. 6).[10] First, education level, or institutionalised cultural capital, follows a roughly south-westerly trajectory: lower cultural capital is associated with authoritarian views and higher cultural capital is associated with liberal views. Coordinates for all the presented categories are important on one or both axes, according to test values, with the exception of GNVQs/5+GCSEs (these are mid-level secondary qualifications), which sit close to the centre of the space (Table 3). Tertiary education is, however, more strongly associated with Axis 1 than Axis 2, as indicated by not just coordinates but the relative size of test values, suggesting that its holders are typically more liberal-left than liberal-right in ethos. This is in contrast to holders of A levels – higher-level secondary qualifications – who are, by the same metrics, strongly associated with Axis 2 but situated more centrally on Axis 1. Student's t-tests suggest that, on Axis 2, the distances between, on the one hand, the mean for those with A levels and, on the other hand, the means for those with PhDs ($t=2.59$, $df=173$, $p<0.01$), Masters degrees ($t=2.47$, $df=278$, $p<0.05$)

[10] The schema is based on the ISCO 08 variable in the dataset. All supplementary variables are treated as multiple nominal.

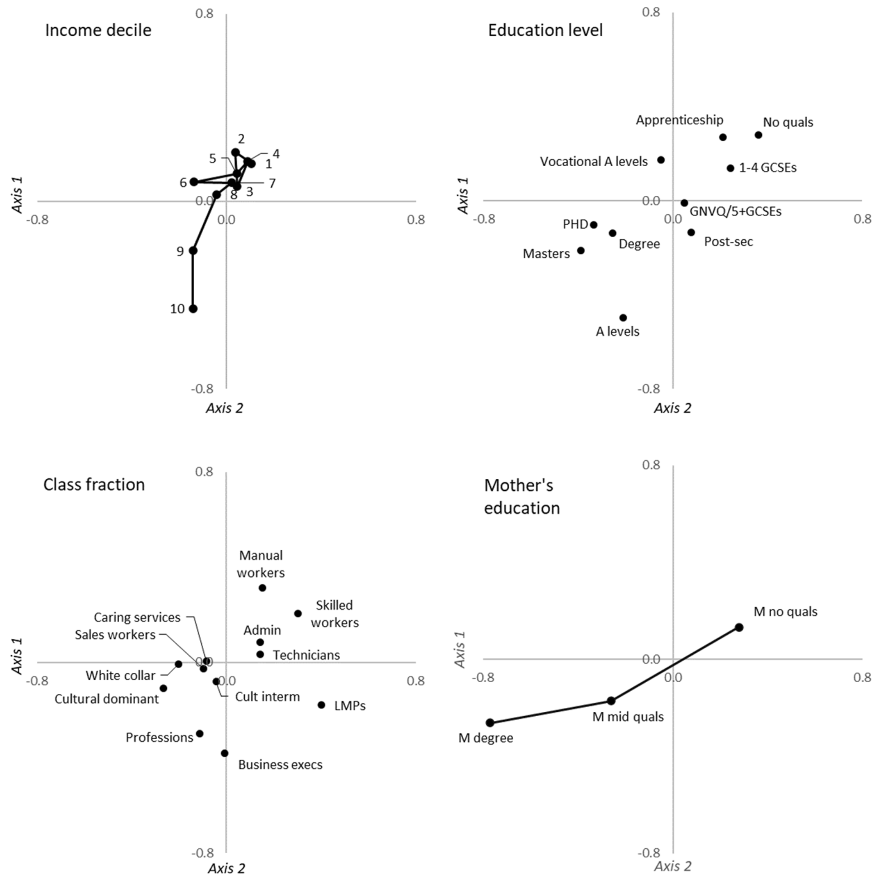

Fig. 6 Capital and class position in the CatPCA model (category means)

and undergraduate degrees ($t=3.74$, $df=424$, $p<0.001$) are important, even if effect sizes are in the small-to-medium range ($d=0.44$, 0.28, 0.37 respectively).

Mother's education – as an indicator of 'inherited' cultural capital – is also strongly related to the axes. In this instance, though, the relationship with the first axis is much stronger than with the second axis: test values are noticeably larger on Axis 1 and the F statistic and eta-squared figure are three times those for Axis 2. The distance between mean points for mothers with no education and degree educated mothers is significant on axis 2 ($t=5.92$, $df=1072$, $p<0.001$) but the relationship is much stronger on axis 1 ($t=9.74$, $df=1072$, $p<0.001$). This is reflected in the respective Cohen's d figures: on Axis 2 it is 0.49 – a medium effect – but on axis 1 it is 0.88, and thus comfortably large.

Second, the relationship between the second axis and economic capital is plain: although there is little differentiation between lower and middling categories, the top two deciles plunge southwards in the space. The means of the bottom and top two categories are important on the second axis, as indicated by test values, and the variable as a whole is, according to analysis of variance, substantially dispersed on Axis 2 – but *not* on Axis 1. There is,

Table 3 Test statistics for supplementary variables in the CatPCA space

	Axis 1	Axis 2		Axis 1	Axis 2
Income Decile			*Education*		
1	1.31	2.08*	PhD	-2.72**	-0.83
2	0.36	2.67**	Masters	-5.34**	-2.90**
3	0.50	0.79	Degree	-4.97**	-2.66**
4	0.95	1.93	Post-sec	1.31	-2.59**
5	0.44	1.43	Vocational	-0.80	2.23*
6	-1.74	0.93	Apprenticeship	2.71**	3.52**
7	0.17	0.99	A levels	-2.32*	-5.25**
8	-0.59	0.31	GNVQ/5+GCSEs	0.53	-0.18
9	-1.94	-2.75**	1–4 GCSEs	2.17*	1.24
10	-2.30*	-6.80**	No quals	7.19**	5.57**
F	1.54	7.67**	F	13.47**	10.42**
Eta^2	0.01	0.04	Eta^2	0.06	0.05
Class fraction			*Mother's education*		
Business execs	-0.13	-3.31**	M no quals	10.50**	4.94**
Professions	-0.97	-2.43*	M mid quals	-7.51**	-4.86**
White collar	-2.83**	-0.14	M degree	-7.98**	-4.97**
Cultural dominant	-3.39**	-1.38	F	85.80**	26.65**
LMPs	3.71**	-1.70	Eta^2	0.10	0.03
Technicians	1.31	0.29			
Administrators	1.76	1.08			
Cultural interm	-0.47	-0.77			
Skilled workers	4.05**	2.74**			
Manual workers	2.41*	5.30**			
Sales workers	-1.53	-0.45			
Caring services	-1.23	0.04			
F	5.49**	5.03**			
Eta^2	0.03	0.03			

* $p<0.05$, **$p<0.01$.

then, some differentiation in direction of travel between economic and cultural capital at the higher levels, as if to signify the importance of capital composition in structuring political attitudes. In short, higher cultural capital is strongly associated with liberalism and a broadly centre-left view on material matters whereas higher economic capital – and only middling acquired cultural capital – is associated with liberal-right views.

Confirming the patterns so far, but alluding to homology with the broader structure of the British social space, is the distribution of the categories of the Bourdieusian class scheme. There are, for sure, some categories clustering around the barycentre, which suggests a tendency toward 'the middle ground' among those class fractions – this is what the MCA detected – but more categories possess coordinates in the space that roughly approximate the expected structure of the homology from previous research (esp. Atkinson 2017): the cultural dominant are associated with liberal/centre-left views, business executives with liberal-right views and the white-collar workers in between; manual workers, skilled workers and technicians tend toward authoritarian views but are differentiated by their gravitation toward left or right-wing views on material politics; and the LMPs are associated above all with right-wing views on materialist politics.

Test values on the axes reflect these relationships, including the differentiation of the cultural dominant from business executives within the dominant class: the former are associated with Axis 1 whereas business executives are not, and the relationship reverses on Axis 2. Distances between extremity categories, as given by t-tests, are also important. If the cultural dominant is taken as the reference group, then the most substantial distances on Axis 1 are from skilled workers ($t=-4.89$, $df=327$, $p<0.001$, $d=0.55$), LMPs ($t=-4.85$, $df=236$, $p<0.001$, $d=0.62$) and manual workers ($t=-3.86$, $df=426$, $p<0.001$, $d=0.42$). On Axis 2, the most notable distances are between business executives on the one hand and manual workers ($t=-5.23$, $df=353$, $p<0.001$, $d=0.66$) and skilled workers ($t=-4.33$, $df=254$, $p<0.001$, $d=0.61$) on the other, but the gap between business executives and the cultural dominant is also important, if more modest in size ($t=1.97$, $df=223$, $p<0.05$, $d=0.28$). All in all, then, there is everything to say that class definitely matters as a structuring force of the space, and *not only in terms of capital volume but capital composition* – a finding that now concords with previous research.

7 Conclusion

One could examine further properties of the space, and of course one could explore the effects of category aggregations or selecting different or further active variables to construct the space. But that has not been the point. Nor have I aimed to reject everything that Lindell and Ibrahim argued: I confirmed the first dimension of their model, polarising left-liberalism and right-authoritarianism, and I have not disputed that intensity of opinion is a notable feature of political attitudes. Instead I have used a different method, suggested by the data itself, to reveal what was hidden: the polarisation of right-wing liberalism and left-authoritarianism and the multidimensionality of the relationship with class. In doing so I hope to encourage greater methodological flexibility and adaptability among those wanting to construct models of fields, social spaces and symbolic spaces in Bourdieusian fashion – to make clear that MCA is not the only or necessarily the best way to chart a field or space and can be complemented by other methods. More specifically, I recommend the incorporation of CatPCA into the Bourdieusian GDA fold, from which it is currently absent, because it offers an efficient means of counteracting horseshoe effects where they prove stubborn. It requires thinking in terms of vectors as well as categories, but it nonetheless produces relational structures of association, opposition, proximity and distance. To repeat, this is not to suggest CatPCA *replace* MCA, even when faced with ordinal data only, but rather that CatPCA complement it: MCA may be the default method, but if it turns up powerful and dogged horseshoes even after recoding then CatPCA can come into the fray, telling us what structures lie within the data when its ordinality is accounted for and allowing us to explore the properties of the space robustly. It would even be good practice to report this process and acknowledge the power of the middle/extreme or intensity dimension before proceeding to describe the CatPCA operations and results, so that nothing is hidden.

With greater methodological flexibility may come greater epistemological vigilance, that is, greater reflection on the relationship between statistical models and the social structures they are held to provide approximate images of. It is imperative that the output from an MCA or any other method not be taken too readily as a reflection of what social relations 'really' look like. Models are dependent on the variables entered, for one thing, which might be of greater or lesser suitability to the task at hand or thesis being investigated, and thus yield a more or less distorted image. Explained inertia in a model should also not be too readily equated with *structural* salience and *sociological* importance since it can fluctuate according to all manner of factors, including, as we have seen, which method one adopts. This is not to stubbornly shield theoretical constructs – like the capital composition principle – from empirical refutation but to insist that theory precede and guide method, including exploration of multiple measures and techniques to determine which are best suited to doing the constructs justice with the data at hand, even if one must always remain open to the possibility that the social world is telling us something different from what we originally expected.

Funding The author declares that no funds, grants, or other support were received during the preparation of this manuscript. There are no relevant financial or non-financial interests to disclose.

Ethics declarations The author declares no conflict of interest.

References

Atkinson, W.: Class in the New Millennium. Routledge, London (2017)
Atkinson, W.: The Class Structure of Capitalist Societies, Volume 1: A Space of Bounded Variety. Routledge, London (2020)
Atkinson, W.: The Swedish Space of Lifestyles and Symbolic Domination. Social Science Information. **60**(1), 63-85 (2021)
Atkinson, W.: The Class Structure of Capitalist Societies, Volume 2: Social Space and Symbolic Domination in Three Nations. Routledge, London (2022a)
Atkinson, W.: The US Space of Lifestyles and its Homologies. Sociological Perspectives. **65**(6),1060-1080 (2022b)
Atkinson, W.: Finding the Best Tool for the Job: The Case of the British Social Space and its Homologies. SPAIS Working Paper 01-22, http://www.bristol.ac.uk/spais/research/workingpapers/ (2022c)
Atkinson, W., Marzec, P.: The German Space of lifestyles: A Multidetermined Structure. Journal of Consumer Culture. doi: 10.1177/14695405221149117 (2023)
Atkinson, W., Rosenlund, L.: Mapping the British Social Space. SPAIS Working Paper 02-14, http://www.bristol.ac.uk/spais/research/workingpapers/ (2014)
Barth, A., Schmitz, A.: Response quality and ideological dispositions. Qual. Quantity. **52**, 175–194 (2018)
Benzecri, J.-P.: Correspondence Analysis Handbook. Dekker, New York (1992)
Blasius, J., Greenacre, M. (eds.): Visualization and Verbalization of Data. CRC Press, London (2014)
Blasius, J., Müchlichen, A.: Identifying audience segments applying the "Social Space Approach". Poetics. **38**(1), 69–89 (2010)
Blasius, J., Lebaron, F., Le Roux, B., Schmitz, A. (eds.): Empirical Investigations of Social Space. Springer, Cham (2020)
Bourdieu, P.: Distinction. Routledge, London (1984)
Bourdieu, P.: Homo Academicus. Polity, Cambridge (1988)
Bourdieu, P.: The State Nobility. Polity, Cambridge (1996)
Bourdieu, P.: A Conservative Revolution in Publishing. Translation Stud. **1**(2), 123–153 (2008)
Bourdieu, P., Wacquant, L.: An Invitation to Reflexive Sociology. Polity, Cambridge (1992)

Bourdieu, P., Chamboredon, J.-C., Passeron, J.-C.: The Craft of Sociology. Walter de Gruyter, New York (1991)
Bry, X., Robette, N., Roueff, O.: A dialogue of the deaf in the statistical theater? Qual. Quantity. **50**, 1009–1020 (2015)
Cohen, J.: Statistical Power in the Behavioural Sciences, 2nd edn. Lawrence Erbaum, New York (1988)
Comrey, A.: A First Course on Factor Analysis. Academic Press, London (1973)
Coulangeon, P., Duval, J. (eds.): The Routledge Companion to Bourdieu's Distinction. Routledge, London (2015)
Di Franco, G.: Multiple correspondence analysis: One only or several techniques? Qual. Quantity. **50**(3), 1299–1315 (2016)
Doolan, K., Tonkovic, Z.: Convergences and Divergences of a Bourdieusian and occupation-based Approach to Social Class. ' European Societies (2021)
Flemmen, M.: The politics of the service class. Eur. Soc. **16**(4), 543–569 (2014)
Flemmen, M., Haakestaad, H.: Class and politics in twenty-first century Norway. Eur. Soc. **20**(3), 401–423 (2018)
Flemmen, M., Jarness, V., Rosenlund, L.: Social space and cultural class divisions. Br. J. Sociol. **69**(1), 124–153 (2018)
Flemmen, M., Jarness, V., Rosenlund, L.: Class and Status. Br. J. Sociol. **70**(3), 816–866 (2019)
Gifi, A.: Nonlinear Multivariate Analysis. Wiley, New York (1990)
Greenacre, M.: Correspondence Analysis in Practice, 3rd edn. Chapman and Hall, Boca Raton (2017)
Greenacre, M., Blasius, J. (eds.): Multiple Correspondence Analysis and Related Methods. Taylor and Francis, London (2006)
Grenfell, M., Lebaron, F. (eds.): Bourdieu and Data Analysis. Peter Lang, Oxford (2013)
Hanquinet, L., Savage, M. (eds.): Routledge International Handbook of the Sociology of Art and Culture. Routledge, London (2018)
Harrits, G., Prier, A., Rosenlund, L., Skjott-Larsen, J.: Class and politics in Denmark. Scandinavian Political Studies. **33**(1), 1–27 (2010)
Hjellbrekke, J.: Multiple Correspondence Analysis for the Social Sciences. Sage, London (2019)
Husson, F., Le, S., Pagès, J.: Exploratory Multivariate Analysis by Example Using R. CRC Press, London (2017)
Inglehart, R.: The Silent Revolution. Princeton University Press, Princeton (1977)
Inglehart, R.: Culture Shift in Advanced Industrial Society. Princeton University Press, Princeton (1990)
Inglehart, R.: Modernization and Postmodernization. Princeton University Press, Princeton (1997)
Jarness, V., Flemmen, M., Rosenlund, L.: From Class Politics to Classed Politics. Sociology. **53**(5), 879–899 (2019)
Joye, D., Birkelund, E.G., Lemel, Y.: 'Travelling with Albert Gifi' in Blasius, J., Lebaron, F., Le Roux, B. and Schmitz, A. (Eds) (2020) Empirical Investigations of Social Space. Springer, Cham, pp. 393–410. (2020)
Lamont, M.: Money, Morals and Manners. University of Chicago Press, Chicago (1992)
Le Roux, B.: Analyse Géometrique des Données Multidimensionnelles. Dunod, Paris (2014)
Le Roux, B., Rouanet, H.: Geometric Data Analysis. Kluwer, Dordrecht (2004)
Le Roux, B., Rouanet, H.: Multiple Correspondence Analysis. Sage, London (2010)
Le Roux, B., Rouanet, H., Savage, M., Warde, A.: Class and Cultural Division in the UK. Sociology. **42**(6), 1049–1071 (2008)
Le Roux, B., Bienaise, S., Durand, J.-L.: Combinatorial Inference in Geometric Data Analysis. CRC Press, London (2020)
Lebart, L., Morineau, A., Warwick, K.: Multivariate Descriptive Statistical Analysis. Wiley and Sons, New York (1984)
Lebart, L., Piron, M., Morineau, A.: Statistique Exploratoire Multidimensionelle. Dunod, Paris (2006)
Lemel, Y., Katz-Gerro, T.: The stratification of leisure. Loisir et Société/Society and Leisure. **38**(3), 399–342 (2015)
Lindell, J., Ibrahim, J.: Something 'Old', something 'New'? The UK Space of Political Attitudes after the Brexit Referendum. Sociol. Res. Online. (2020). https://doi.org/10.1177/1360780420965982
Maire, Q.: Moral cosmopolitanism, civic action and ethical consumption. J. Sociol. (2021). https://doi.org/10.1177/14407833211053897
Majima, S., Savage, M.: Have there been culture shifts in Britain? Cult. Sociol. **1**(3), 293–315 (2007)
Marzec, P.: Distinction in Poland. Unpublished PhD thesis, University of Bristol. (2019)
Meuleman, R., Savage, M.: A field analysis of cosmopolitan taste. Cult. Sociol. **7**(2), 230–256 (2013)
Norris, P., Inglehart, R.: Cultural Backlash. Cambridge University Press, Cambridge (2019)
Prieur, A., Rosenlund, L., Skjott-Larsen, J.: 'Cultural Capital Today' Poetics. **36**(1), 45–71 (2008)
Robson, K., Sanders, C. (eds.): Quantifying Theory: Pierre Bourdieu. Springer, Dordrecht (2009)

Rosenlund, L.: Exploring the City with Bourdieu. Dr Muller, Saarbrucken (2009)
Rosenlund, L.: ‘Working with *Distinction*’ in Coulangeon, P. and Duval, J. (Eds) (2015) The Routledge Companion to Bourdieu’s Distinction London: Routledge, pp. 157–86. (2015)
Savage, M.: Identities and Social Change in Britain Since 1940. Oxford University Press, Oxford (2010)
Savage, M.: The Return of Inequality. Stanford University Press, Stanford (2021)
Schmitz, A.: The Structure of Digital Partner Choice. Springer, Cham (2017)