Correcting Acquiescence Bias in the Case of Fully Unbalanced Scales with Application to UK Measurements of Political Beliefs

Phil Swatton

Abstract

A substantive area of interest in public opinion research is the political beliefs of voters. Yet, measuring such beliefs can be difficult, specifically when researchers rely on ‘off the shelf’ datasets. Many of these datasets contain unbalanced Likert scales, which risk acquiescence bias. This paper proposes a strategy for dealing with this issue. I first demonstrate using two comparable datasets from the UK how unbalanced scales produce distorted distributions and can affect regression results. Then, building on past research that utilises factor analysis to eliminate the influence of acquiescence bias, I demonstrate how researches can utilise a random-intercepts confirmatory factor analysis method to obtain factor scores corrected for acquiescence in the case of fully unbalanced scales. The paper concludes with practical recommendations for researchers and survey designers moving forward.

Paper prepared for presentation at APSA Conference, 1 - 3 Oct 2021

Early draft. Please do not cite, quote, or circulate without the authors’ permission.

Introduction

In public opinion research, a substantive area of interest is the political beliefs of voters. It follows that for such research to be conducted, adequate measurements of these beliefs must be obtained. In practice, due to the expense and time involved in survey research public opinion researchers will often prefer or need to use ‘off the shelf’ survey data produced by others rather than conduct their own surveys. A commonly used method in these datasets for measuring beliefs and other variables of interest is the Likert scale. In this question format, respondents are shown statements and given a range of responses, usually five ranging from ‘strongly disagree’ to ‘strongly agree.’ Scores from these responses are then tallied to produce a final measurement of the concept of interest. However, many popularly used public opinion datasets contain ‘unbalanced’ rather than ‘balanced Likert scales. Consequentially, they suffer from acquiescence bias. While the effect of acquiescence bias on Likert scales has been known for some time, many popular ’off the shelf’ datasets continue to contain either unbalanced scales or even agree-disagree items on their own. If researchers utilising these surveys fail to consider the survey design and response stage of the data generating process (DGP), this will likely lead to incorrect research conclusions being drawn.

I begin this paper with a brief discussion of acquiescence bias in the context of measurement theory. From here, I proceed to demonstrate how acquiescence bias produces different results from unbalanced to balanced scales by utilising data from the 2017 British Social Attitudes (BSA) survey (NatCen-Social-Research 2017) and the 2017 British Election Study (BES) face-to-face post-election survey (Fieldhouse et al. 2017). These surveys were collected in roughly the same time period and crucially have identical theoretical motivations for the scales I examine. The scales in these surveys differ in a key respect: the BES survey contains (mostly) balanced scales, while the BSA contains fully unbalanced scales. Consequentially, any substantive differences in correlations, distributions, and regression results should not stem from either temporal differences in public opinion or conceptual differences underlying chosen measurements. I therefore use these scales to show the differences that emerge between the distributions of the measurements, the effect of education on the measurements, and to establish a baseline difference in results for attempting any correction.

Past research has offered a useful method for correct acquiescence bias via unit-intercept confirmatory factor analysis (CFA)[1] to control acquiescence bias in the case of balanced and partially unbalanced scales. However, the more difficult case of fully unbalanced scales has been left unexamined. I therefore modify past research in developing a method for balanced and partially unbalanced scales to develop a method for obtaining acquiescence-corrected factor scores. The modified version of the method requires a further partially unbalanced or balanced scale to successfully identify the acquiescence bias component within the fully unbalanced scale(s) of interest, as the data on its own does not offer information which can be used to distinguish between genuine agreement and acquiescent agreement.

I develop four variations of the unit-intercept CFA model, based on two distinctions. The first is whether the indicators are treated as continuous (CFA) or ordinal (OCFA), and the second is two variations in how the unit intercepts are specified (methods 1 and 2). I then run a Monte Carlo (MC) simulation to assess the ability of the four variations to successfully capture the content factor of interest while removing the acquiescence bias. Unlike past research which for the most part places its emphasis on recovering substantive factor loadings, my emphasis is instead on recovering bias-free content factors as it is this task that is most relevant to applied research. The results of the MC simulation suggest that variation 1 is to be preferred - with some caveats - while explicitly modelling the data as ordinal does not seem to offer much.

I then proceed to examine a real-world application of the proposed methodology to see how use of the four variations might affect substantive research results. To this end I utilise data from Wave 14 of the British Election Study Internet Panel (BESIP) (Fieldhouse et al. 2020) as a cross-sectional dataset. The choice of this dataset is motivated in part due to its containing similarly unbalanced Likert scales as the BSA and partially due to its meeting the data requirements of the method when dealing with fully unbalanced scales. I exploit the presence of two balanced scales in this wave to estimate the four unit-intercept models. I assess how the predicted factor scores differ in their distributions and how substantive results change when regressed on education. Despite the issues highlighted in the simulation study, the OCFA model for version 2 of the unit-intercepts specification does seem to perform best at producing results with a reasonable degree of face validity in matching the BES face to face results earlier in the paper.

While I have developed some approaches for applied researchers in dealing with fully unbalanced scales, the method retains two serious limitations. The first is that it is situational: it can only be performed where there is a balanced or partially unbalanced scale (or more accurately indicators for these scales) available for identification of the acquiescence bias. The second major limitation is that it assumes not only equal item acquiescence across the items for a given scale, but across all scales utilised and analysed. The first limitation is unavoidable in the case of fully unbalanced scales and will necessarily apply to any other methods developed for them. It is clear from the MC simulation results that the second limitation has the potential to be restrictive conditional on the similarity in item acquiescence from the first-step to second-step scales.

I therefore conclude with both recommendations for researchers utilising surveys and researchers designing surveys. For those utilising these surveys, I (re-)highlight the need to be aware of how acquiescence bias may affect substantive research results and through both the MC simulation and substantive results offer some considerations in choosing potential correction methods. Where scales are fully unbalanced and no correction can be performed, the recommendation is a specific case of a more general one: that researchers should consider not only the DGP prior to responses being collected but also the DGP of survey design and responses. In practical terms, this requires that researchers adapt their interpretations of research results to be more cautious. For researchers designing surveys, the primary recommendation is more specific but well-established: to utilise balanced scales. However, to facilitate cross-temporal comparisons, I recommend that designers of 'off the shelf' surveys either add additional indicators without removing any to facilitate balanced scale or ensure there is at least one balanced scale present to facilitate bias correction.

[1]Some recent research has placed this in the IRT framework, but the way it works is identical to the CFA approach.

Acquiescence Bias

In fully unbalanced scales, all questions are asked with the 'strongly agree' response matching the same end of a particular concept scale (e.g. the 'left' end of the economic dimension). In other words, it would be consistent for a respondent to agree with all statements presented (or disagree with all). Conversely, in balanced scales half of all statements match one end of the concept scale of interest and the other half match the other end. Implicitly this requires an even number of statements. Assuming sensible choices have been made for the statements it should be reasonably contradictory to agree with all statements (or similarly, to disagree with all statements). An unbalanced scale is then simply one in which there are more 'agree' items on one end of the concept scale than the other. Balanced scales are preferred because of the existence of a response bias known as 'acquiescence bias.' Put simply, acquiescence bias is the tendency for individuals to be more likely to agree with a statement regardless of its actual content (Mirowsky and Ross 1991; Billiet and McClendon 2000; Billiet and Davidov 2008). In general, acquiescent survey items will be more correlated with each other than they should be, and less correlated with non-acquiescent survey items than they should be (Winkler, Kanouse, and Ware 1982, 557).

This occurs because any given indicator variable will in practice likely contain some variation from the variable of interest, some variation from the measurement method of choice, and some variation unique[2] to that indicator (Kenny and Kashy 1992, 165). Consequentially, when the measurement-related variation is either positively or negatively correlated with the concept of interest, the final measurement's correlation with other acquiescent variables will be higher than they should be. Similarly, the absence of acquiescence-related variation in other substantive variables of interest will result in correlations that are lower than they otherwise should be. Crucially, in the case of Likert scales because acquiescence leads to agreement **regardless** of statement content when all statements are in agreement (or indeed if more statements agree with one end of the concept scale than the other) then the underlying content variation will be correlated with the underlying acquiescence variation.

As implied by the distinction between balanced and unbalanced scales, the standard solution to acquiescence bias occurs at the survey design stage. Where an equal number items are worded (or 'keyed') in opposite directions, it is assumed the acquiescence present in individual indicator is 'cancelled out' in the creation of the final scale (see e.g. Cloud and Vaughan 1970; Ray 1979; G. A. Evans and Heath 1995). This strategy is conditional on the assumption that acquiescence will equally affect responses for all statements in the scale (or at least the summed acquiescence for individuals across all items will be 0) which is often incorrect (Billiet and Davidov 2008, 545). Of course, it may not need to be entirely correct to be reasonably successful - it need only be a reasonable approximation of the degree of acquiescence in each item. In other words, the strategy will presumably be successful despite its strong assumption if the acquiescence across items is reasonably similar, even if not fully identical.

It is therefore clear that where 'off the shelf' datasets contain unbalanced scales, this will lead to biased distributions and biased correlations between measured variables. The danger is however greater again. Acquiescence bias does not occur solely due to the measurement method of choice - respondents in themselves are also a source of variation in acquiescence (Billiet and Davidov 2008, 543). In particular, it is reasonably well-established that respondents with lower income and education levels tend to be more acquiescent (Ware Jr 1978; Winkler, Kanouse, and Ware 1982, 555). Consequentially, not only do measurements of 'latent' variables such as dimensions of belief become negatively affected in the above manner, but their association with 'objective' measures can also become biased. For example, the association of education with an acquiescent measure will become biased against the acquiescent direction of that particular measure as higher levels of education are associated with a lesser tendency towards acquiescence in the respondent.

A Demonstration of the Problem

While the problem of acquiescence bias is well-known, it is useful to begin with a demonstration of its substantive impact. For the purposes of this paper this will serve both as a demonstration of how acquiescence bias affects research and as a baseline against which subsequent correction results can be compared. Two

[2]The unique component is a mixture of noise and substantive variation not in other indicators.

datasets from Great Britain - the British Social Attitudes survey (BSA) and British Election Study (BES) face-to-face survey are chosen. The British case is chosen not because this issue is constrained to any particular location (it will occur wherever unbalanced scales are used, or even indicators on their own), but because many British datasets contain similar scales for political beliefs with the same conceptual basis. While many scholars have worked on the notion of a second dimension of political belief and the notion of multidimensional politics, the most influential in the UK are Evans and Heath (see Heath, Evans, and Martin 1994; G. A. Evans and Heath 1995; G. Evans, Heath, and Lalljee 1996). Most UK datasets with measurements of political beliefs contain likert scales either inspired by or directly derived from these papers. One consequence of the common inspiration for these scales is that where differences in how these scales are implemented occur so too do the best opportunities to analyse the consequences of these differences.

Evans and Heath et al in these papers define two primary dimensions of interest: a traditional economic dimension ranging from left to right and a second dimension of beliefs ranging from libertarian to authoritarian. In this conceptualisation, the former (alternatively labelled socialist vs lassaiz-faire) represents issues of (economic) equality, while the latter represents issues of personal freedom (G. Evans, Heath, and Lalljee 1996). In this paper I will be dealing solely with acquiescence in scales derived from this set of research, due in no small part to their widespread availability across several datasets. For the purposes of demonstration, I will be utilising the 2017 British Social Attitudes (BSA) survey (NatCen-Social-Research 2017) and the 2017 British Election Study (BES) face-to-face survey (Fieldhouse et al. 2017). The BSA survey was conducted from July to October with a small number of surveys in November. The BES survey was meanwhile conducted from June to October 1st. Both contain scales derived from the work of Evans and Heath. However, in the BSA both scales are fully unbalanced, while in the BES the scales are balanced and partially unbalanced.

With near-complete overlap in the time periods for data collection and identical conceptual bases, it is plausible to assume that substantial differences in the distributions (and other any other results) are a consequence of the acquiescence bias implied by the fully unbalanced structure of the BSA. The items used to construct the scales can be reviewed below. In the BSA, all items are left-wing or authoritarian. For the BES, it is noted next to each item whether the statement is left-wing, right-wing, libertarian, or authoritarian in its direction. All of the scales where constructed to range from 0 (left/libertarian) to 4 (right/authoritarian). Since survey weights are used throughout unless otherwise stated, the BES respondents without survey weights were not included all parts of the following analysis.

BSA

The statements utilised in the BSA economic dimension (ranging from Disagree Strongly to Agree Strongly) are as follows:

- Government should redistribute income from the better off to those who are less well off
- Big business benefits owners at the expense of workers
- Ordinary working people do not get their fair share of the nation's wealth
- There is one law for the rich and one for the poor
- Management will always try to get the better of employees if it gets the chance

The statements utilised in the BSA second dimension (ranging from Disagree Strongly to Agree Strongly) are as follows:

- Young people today don't have enough respect for traditional British values
- People who break the law should be given stiffer sentences
- For some crimes, the death penalty is the most appropriate sentence
- Schools should teach children to obey authority
- The law should always be obeyed, even if a particular law is wrong
- Censorship of films and magazines is necessary to uphold moral standards

BES

The statements utilised in the BES economic dimension (ranging from Strongly Disagree to Strongly Agree) are as follows:

- Ordinary working people get their fair share of the nation's wealth (right)
- There is one law for the rich and one for the poor (left)
- There is no need for strong trade unions to protect employees' working conditions and wages (right)
- Private enterprise is the best way to solve Britain's economic problems (right)
- Major public services and industries ought to be in state ownership (left)
- It is the government's responsibility to provide a job for everyone who wants one (left)

The statements utilised in the BES second dimension (ranging from Strongly Disagree to Strongly Agree) are as follows:

- Young people today don't have enough respect for traditional British values (auth)
- Censorship of films and magazines is necessary to uphold moral standards (auth)
- People should be allowed to organise public meetings to protest against the government (lib)
- People in Britain should be more tolerant of those who lead unconventional lives (lib)
- For some crimes, the death penalty is the most appropriate sentence (auth)
- People who break the law should be given stiffer sentences (auth)

Expected Consequences

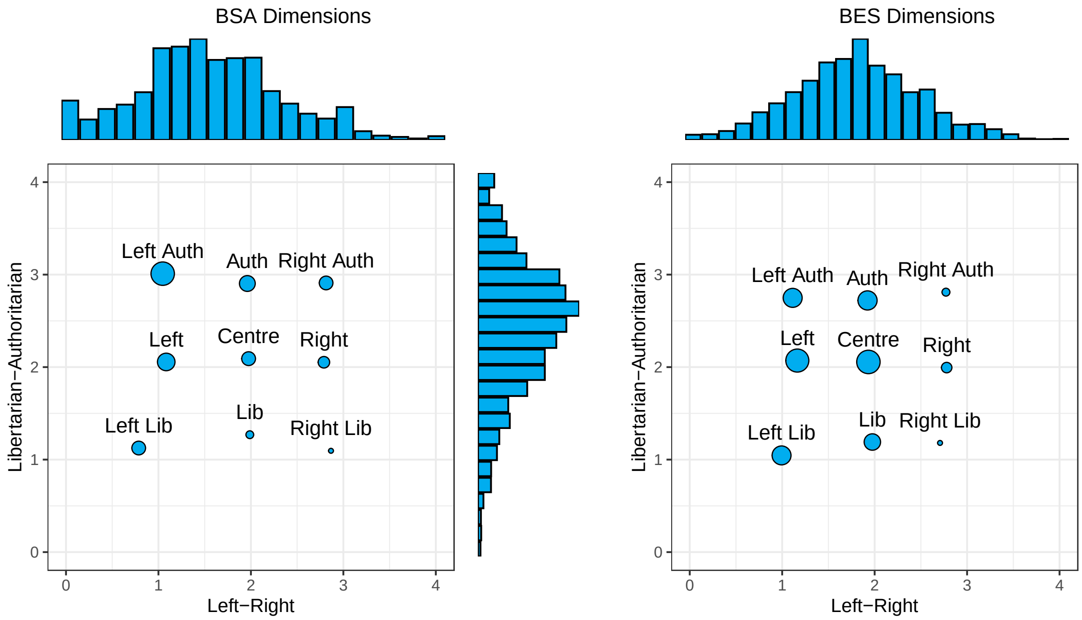
Before proceeding, it is useful to explain what differences in results should be expected given the preceding discussion. Due to the fact that in the BSA all statements are left-wing or authoritarian, the direction of acquiescence bias should be in these directions. In other words, relative to the (mostly) balanced scales in the BES, respondents in the BSA should be more left-wing and authoritarian. Assuming the true correlation between right-wing attitudes and authoritarian attitudes is either null (i.e. 0) or positive (i.e. >0), acquiescence bias will result in a correlation between the scales that is moved in a negative direction, due to the fact that the acquiescent component in the BSA left-right and libertarian-authoritarian scale will be pointing in opposite directions (moving towards 0 in the former, moving towards 4 in the latter). Finally, it is reasonably well-established that education has a negative association with authoritarianism but there is no well-established result suggesting it has a similar relationship with right-wing attitudes. We can therefore predict that when regressing these scales on an education variable, there will be a positive association between the BSA left-right scale and higher levels of education (as acquiescence goes in the left-wing direction), while the BES left-right scale will likely see a null result for the effect of education. In both libertarian-authoritarian scales there will be a negative association between education and the scales, but it will be larger in the BSA scales due to the fact that it is mixed up with acquiescence bias.

Difference in Distributions

To demonstrate the impact of acquiescence bias on the distributions of these scales, figure 1 presents both the joint and marginal distributions of these scales in two graphs. The left-hand graph presents the distributions for the BSA scales while the right-hand graph presents the distributions for the BES scales. As described above, the scales where constructed from the indicators in both surveys and constructed to range from 0 to 4. In both plots, the x-axis represents the left-right scale while the y-axis represents the libertarian-authoritarian scale. The corresponding histograms are the histograms for those respective scales, showing their marginal distributions. To visualise the join distributions of these scales, respondents were divided into 'groups' for each of the axes. Those with scores ranging from 0 to 1.6 were placed in the 'left' and 'libertarian' groups of the respective dimensions. Correspondingly, those with scores ranging from 2.4 to 4 were placed in the 'right' and 'authoritarian' groups of the respective dimensions. Finally, those in-between these values were placed in the 'centre' group for each dimension. These groupings are of course arbitrary, but were chosen in part to resemble similar groupings utilised in research using these scales (see Surridge 2018) as a means of emphasising how consequential acquiescence bias can be for research conclusions. The mean of each group was plotted, while the size of the group's dot corresponds to the number of respondents in that group. Survey weights were used for the graphs.

The graphs in figure 1 show some clear differences between the distributions in the BSA and BES surveys. First, as expected, the BSA shows a clear left and authoritarian slant as compared to the BES. Interestingly, there is a larger difference on the libertarian-authoritarian scale than on the left-right one, despite the fact it

Figure 1: Joint Distribution of BSA and BES Scales



is the latter that remains partially unbalanced. The BES graph suggests there is still a somewhat left-leaning slant in the British electorate, similar to the BSA. This is however driven by the relative lack of right-wing respondents, as there has still been an overall shift towards the center. The weighted mean for the BSA left-right scale is 1.52, while the BES equivalent is 1.63 - suggesting a small difference albeit one in line with expectations. To check whether this difference was meaningful, I ran a weighted t-test. This produced a p-value lower than 0.05 and thus I reject the null hypothesis of the difference between the means being 0. The distribution of the libertarian-authoritarian scale in the BES is considerably more balanced again, with a fairly even spread across the dimensions. The weighted mean of the BSA libertarian-authoritarian scale is 2.47 while the BES equivalent is 2.04. I performed a weighted t-test on the differences in these scales and once again the p-value was below 0.05 and thus I rejected the null hypothesis of no difference.

Broadly, the differences here are in line with the expectations I set out in the earlier portion of this section. The BES left-right and libertarian-authoritarian scales are more balanced than their BSA counter-parts; and the direction of the differences between them are in line with those predicted by some knowledge of acquiescence bias. There has been one surprise in the form of relatively few right-wing respondents in the BES. While the BES data therefore would offer firmer grounds for believing that the British electorate in 2017 was left-wing, some caution is still required. First, this may of course simply be a quirk of the sample. It may also be a function of the statements provided - if for instance some of the items are only weakly associated with the underlying left-right views of the electorate and most people agree with them regardless of ideological difference, this could skew the results when Likert scales are used. This is because Likert scales assume that all items in the scale equally capture the content of interest, rather than allowing for the fact that the balance of content and unique variation will change from indicator to indicator.

Difference in Associations

With the difference in distributional results demonstrated, I now turn to demonstrating the difference in associations between variables. My interest here is not to offer some causal explanation of these scales, but rather to offer a clear example of results changing from unbalanced to balanced scales. I therefore briefly discuss the between-scale correlations, before regressing each scale on education. The education variables[3] in both surveys were recoded such that the categories would match - see appendix tables A1 and A2 for the

[3] HEdQual2 in the BSA, edlevel in the BES.

BSA and BES respectively. Most of the recodes will be uncontroversial, but the 'foreign' category in the BSA had to be treated as missing as it had no clear placement. The weighted correlation between these scales in the BSA is 0.03 while for the BES the value is 0.08. The difference is fairly small, suggesting that between-scale correlations may be fairly robust. An alternative form of association we may be interested in is between the scales - either as dependent or independent variables - and other more 'objectively' measured variables. One such example is education level, which as above is known to be a determinant of acquiescence and of some political beliefs. Table 1 shows the results of these regressions. The scales are labelled in the column headings, and the reference category is no educational qualifications.

Table 1: BSA and BES Scales Regressed on Education

	BSA Left-Right	BES Left-Right	BSA Lib-Auth	BES Lib-Auth
Intercept	1.31^{***}	1.65^{***}	2.82^{***}	2.26^{***}
	(0.04)	(0.02)	(0.03)	(0.02)
GCSE/Equiv	0.22^{***}	0.01	-0.23^{***}	0.09
	(0.04)	(0.07)	(0.04)	(0.07)
A-level/Equiv	0.30^{***}	-0.12^{**}	-0.32^{***}	-0.23^{***}
	(0.06)	(0.04)	(0.05)	(0.04)
Undergrad	0.27^{***}	0.02	-0.68^{***}	-0.40^{***}
	(0.05)	(0.04)	(0.04)	(0.03)
Postgrad	0.24^{***}	-0.00	-0.84^{***}	-0.70^{***}
	(0.06)	(0.05)	(0.05)	(0.05)
R^2	0.01	0.01	0.13	0.12
Adj. R^2	0.01	0.00	0.13	0.12
Num. obs.	3123	1806	3125	1931

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

In line with the expectations outlined above, there are two notable shifts in the results from the BSA to the BES. First, and least dramatically, the absolute size of the point estimates of increasing levels of education on libertarian-authoritarian beliefs all decrease. Second, and more dramatically, the positive association between increasing levels of education and right-wing beliefs captured in the BSA scale entirely disappears in the BES. As stated above, this is not a causal analysis and should not be interpreted as one. What it *does* show however, is that acquiescence can represent a potentially drastic confounding variable in regression analysis - and therefore also in causal inference. As with the differences in distributions and correlations, these results are fully in line with what we should expect given some theoretical understanding of acquiescence bias.

All of the results demonstrated in this section have relied on an assumption that there should be no *predictable* differences from the BES to the BSA *other than those caused by acquiescence*. Given the importance of this assumption to my analysis above, I have performed two robustness checks to verify that these differences are likely driven by acquiescence bias, rather than any particular quirks of the samples. First, I merged the five indicators common to the BES and BSA into a single dataset and created a binary variable denoting whether a respondent belonged to the BSA. I then regressed this binary variable on the five common indicators, and ran an OLS, Logit, and Probit model to check against model dependency. The regression table is available in the appendix as table A3. In all three cases, two indicators were significant (For some crimes, the death penalty is the most appropriate sentence; People who break the law should be given stiffer sentences). However, their point estimates pointed in opposite directions, strongly suggesting that these were sample quirks most likely related to the unique componenets of these indicators. Certainly, it did not offer evidence against any of the above interpretation. Next, to verify that there was no temporal instability in results, I regressed the scales from the BES on the survey month of the respondents. The regression table is available in the appendix as table A4. The result showed no association between interview month and scale score. I did not do the same for the BSA as the interview date is not included in the publicly available version of the dataset. Taken together, these two checks offer strong evidence that my assumption that differences between these two surveys are primarily driven by acquiescence bias is correct.

Unit Intercept Confirmatory Factor Analysis

With the problem of acquiescence bias fully conceptualised and demonstrated, I now turn to the task of developing a method for applied researchers to deal with the problem in the case of fully unbalanced scales. While it remains true that acquiescence bias is best dealt with at the survey design stage, this will be of little comfort to researchers seeking to utilise data collected by other researchers. Papers reviewing competing methodologies for modelling acquiescence bias have concluded that one of the most effective is an approach that treats acquiescence as a person-specific intercept across the scale items (Savalei and Falk 2014; Primi, Santos, et al. 2019; Primi, Hauck-Filho, et al. 2019). This model was developed in a Confirmatory Factor Analysis (CFA)/Structural Equation Modelling (SEM) context (Mirowsky and Ross 1991; Billiet and McClendon 2000) but later extended to a unidimensional Item Response Theory (IRT) context (Primi, Santos, et al. 2019; Primi, Hauck-Filho, et al. 2019). For the sake of simplicity I focus on the CFA specification in this paper, but the general intuition translates to the (multidimensional) IRT context without much difficulty (see Primi, Santos, et al. 2019; Primi, Hauck-Filho, et al. 2019).

Confirmatory Factor Analysis

The standard confirmatory factor analysis model can be given in its linear form as (Brown 2015; Maydeu-Olivares and Coffman 2006):

$$x_{ij} = \lambda_{j1}\eta_{i1} + ... + \lambda_{jm}\eta_{im} + \epsilon_{ij} \tag{1}$$

Where x_{ij} is the jth indicator variable for respondent i, λ_{jm} is the loading of factor m on indicator j, η_{im} is the value of factor m for respondent i, and ϵ_{ij} is the unique component for indicator j for respondent i. This, in essence, is the *common factor model.* The factors represent the common (or measurement) variation in the indicators (and necessarily load on more than one indicator) while the unique component is just that - the variance component unique to indicator j for respondent i. The assumptions of this model are:

1. The means of the common factors are 0
2. The common factors are normally distributed
3. The means of the unique components are 0
4. The unique components are normally distributed
5. The unique components are uncorrelated with the common factors
6. The unique components are uncorrelated with each other

The model can be converted to a more compact matrix form (Brown 2015; Maydeu-Olivares and Coffman 2006):

$$\mathbf{x} = \mathbf{\Lambda}\boldsymbol{\eta} + \boldsymbol{\epsilon} \tag{2}$$

Where $\mathbf{x}$ is the $p \times 1$ vector of indicators, $\mathbf{\Lambda}$ is the $p \times m$ matrix of factor loadings, $\boldsymbol{\eta}$ is the $m \times 1$ vector of factor scores, and $\boldsymbol{\epsilon}$ is the $p \times 1$ vector of unique components. In turn, we can further express the model in terms of covariance matrices:

$$\mathbf{\Sigma} = \mathbf{\Lambda}\mathbf{\Psi}\mathbf{\Lambda}' + \mathbf{\Theta}_{\boldsymbol{\epsilon}} \tag{3}$$

Where $\mathbf{\Sigma}$ is the $p \times p$ variance-covariance matrix of the indicators, $\boldsymbol{\psi}$ is the $m \times m$ variance-covariance matrix of the common factors, and $\mathbf{\Theta}_{\boldsymbol{\epsilon}}$ is the $p \times p$ variance-covariance matrix of unique components which by assumption 6 is a diagonal matrix. When estimated with maximum likelihood (ML), assuming no (further) restrictions are placed on the latent variables means the discrepancy function minimised is (Brown 2015):

$$F_{ML} = ln|\mathbf{S}| - ln|\mathbf{\Sigma}| + trace(\boldsymbol{S}\mathbf{\Sigma}^{\mathbf{-1}}) - p \tag{4}$$

Where $\mathbf{S}$ is the model-implied variance-covariance matrix and p is the number of indicators.

Introducing the Unit Intercept

First used in Mirowsky and Ross' paper *Eliminating Defense and Agreement Bias from Measures of the Sense of Control: A 2 x 2 Index* (1991), the best exposition of the unit-intercept model is in Maydeu-Olivares and Coffman's paper *Random Intercept Item Factor Analysis* (Maydeu-Olivares and Coffman 2006). Specifically developed to deal with the problem of acquiescence bias but potentially generalisable to other forms of differential person functioning (DPF), the unit-intercept model conceptualises acquiescence bias as an intercept which varies by individual but is constant for that individual across all indicators. It is traditionally estimated by modelling the additional unit-intercept as a common factor, but constraining the the loading to 1. In the case where our indicators have a single common factor (which is the case in the example of this paper), the linear form can be given as (Maydeu-Olivares and Coffman 2006):

$$x_{ij} = \lambda_{jc}\eta_{ic} + 1\eta_{ia} + \epsilon_{ij} \tag{5}$$

Where factor c would be the common factor and factor a would be the unit-intercept factor. Maydeu-Olivares and Coffman introduce three further assumptions for this model, which are worth discussing here. The first two are (Maydeu-Olivares and Coffman 2006):

7. The mean of the unit-intercepts is 0
8. The unit intercepts are uncorrelated with the unique components

Thus far, these are simply assumptions 1 and 5 repackaged for treating the unit-intercept factor separately. I therefore do retain them - but this is because the assumptions have already been made. However, Maydeu-Olivares and Coffman make a further assumption:

9. The unit intercepts are uncorrelated with the common factor(s)

This assumption is explained in part by Maydeu-Olivares and COffman's choice of language for the model. They specifically refer to the model as a *random-intercept* model and clearly are aiming to draw a parallel with multilevel regression modelling in their description of the unit-intercept confirmatory factor analysis model (indeed, their formulae reflect this too). However, the comparison is not necessary and more importantly harms the usefulness of the model. In a random-intercepts regression model, the random intercepts are estimated as their own error component (hence the assumption of independence between them and the independent variables). The unit-intercept here is not being estimated as an error term - it is being estimated as another common factor. The orthogonality assumption is thus not required for identification purposes (as other assumptions in the model are), but rather is made for the purpose of this comparison. This unnecessarily confuses things and potentially reduces the desirability of the model - in their review, Salvei and Falk suggest more work is required to explore potential relaxations of the orthogonality assumption. This assumption however is unnecessary to being with, and I therefore drop it and utilise the terminology unit-intercept instead of random-intercept.

Finally, to identify the scales of the common factors in the unit-intercept model, the variances of the common factors are constrained to 1 (as opposed to their first loading being constrained to 1). By contrast, the variance of the unit-intercept is freely estimated. The important feature of such a model is that the loading of the unit-intercept factor is constrained across indicators. A method of creating such an intercept while constraining the unit-intercept variance to 1 would simply be to apply equality constraints to the unit-intercept loadings, such that they were equal across all indicators:

$$x_{ij} = \lambda_{jc}\eta_{ic} + \lambda_{a}\eta_{ia} + \epsilon_{ij} \tag{6}$$

The different between (6) and (5) is that instead of a loading of '1' on η_{ia}, there is now a freely estimated loading lacking a 'j' subscript as it is common to all indicators. There is little difference in terms of the actual fact of a unit-intercept here. The advantage of this specification of the model is interpretation - it is immediately clear how acquiescence bias compares to the content factors of interest in its effect on the scales. Although the assumption of equal loadings for the acquiescence component is a strong one, simulations do suggest that the model is robust to violations of this assumption (Savalei and Falk 2014). Moreover, as compared to other acquiescence correction methods the model carries the not-insignificant advantage that it does not require a balanced scale to work. Instead, the unit intercept merely acts to capture *inconsistency* in observed responses - and thus only requires at least one opposite-worded indicator in order to successfully capture acquiescence bias. For the remainder of this paper, I number the first approach to modelling the unit intercept is 1 and the second 2, and for convenience I will henecforth refer to the two models above as CFA1 and CFA2.

Ordinal Confirmatory Factor Analysis

One potential flaw of the unit-intercept CFA model is that it does not fully take into account the ordinal nature of the indicator variables typical for Likert scales. One option is of course to simply use robust maximum likelihood (MLR) estimation, which returns the same point estimates but adjusts standard errors and test statistics for violations of the normality assumption. While potentially workable in many cases, a better option may be to convert the unit-intercept model to the ordinal case where the ordinal structure may perform better. Indeed, a reasonably recent simulation study has concluded in favour of Diagonally Weighted Least Squares (DWLS) estimation of ordinal data (Li 2016). In ordinal CFA, the relationship between the latent variables and the observed categories are assumed to exist via a threshold relationship:

$$\begin{aligned} x^*_{ij} &= \lambda_{j1}\eta_{i1} + ... + \lambda_{jm}\eta_{im} + \epsilon_{ij} \\ x_{ij} &= K \quad if \quad \tau_{jk} < x^*_{ij} < \tau_{jk+1} \end{aligned} \tag{7}$$

Where x^*_{ij} is the latent variable underlying x_{ij}, K is one of the t values x_{ij} can take on, τ_{jk} is the kth threshold for indicator j, $\tau_{j0} = -\infty$ and $\tau_{jt} = \infty$.

Ordinal CFA makes similar assumptions to continuous CFA:

1. The means of the common factors are 0
2. The common factors are normally distributed
3. The means of the unique components are 0
4. The unique components are normally distributed
5. The unique components are uncorrelated with the common factors
6. The unique components are uncorrelated with each other

It follows that x^*_{ij} is normally distributed with mean 0 and the covariance matrix:

$$\mathbf{\Sigma} = \mathbf{\Lambda\Psi\Lambda}' + \mathbf{\Theta}_{\boldsymbol{\epsilon}} \tag{8}$$

To identify the variances of the unique components, we set

$$\mathbf{\Theta}_{\boldsymbol{\epsilon}} = \mathbf{I} - diag(\mathbf{\Lambda\Psi\Lambda}') \tag{9}$$

such that the covariance matrix becomes a correlation matrix $\mathbf{P}$.

Ordinal CFA is often estimated in a three-step procedure. First, the thresholds are estimated alone using maximum likelihood. The thresholds are often estimated by the corresponding percentage of respondents in each category of the ordinal variable (Yang-Wallentin, Jöreskog, and Luo 2010). Second, the polychoric correlation matrix of the observed indicators is estimated via maximum likelihood. Third, assuming no restrictions are placed on the thresholds, a least squares discrepancy function based on the polychoric correlations can be used:

$$F_{LS} = (\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta}))'\boldsymbol{V}(\hat{\boldsymbol{p}} - \boldsymbol{p}(\boldsymbol{\theta})) \tag{10}$$

Where $\hat{p}$ is the polychoric correlation matrix estimated in the second step, $p(\theta)$ is the model-implied correlation matrix, $\boldsymbol{\theta}$ represents the parameters of the model, and $\boldsymbol{V}$ is a weighting matrix. The choice of weighting matrix determines the exact estimation method being used. If $\hat{\boldsymbol{\Gamma}}$ is an estimate of the asymptotic covariance matrix of estimated polychoric correlations, then:

- Weighted Least Squares (WLS): $\boldsymbol{V} = \hat{\boldsymbol{\Gamma}}$
- Diagonally Weighted Least Squares (DWLS): $\mathbf{V} = diag(\hat{\boldsymbol{\Gamma}})^{-1/2}$
- Unweighted Least Squares (ULS): $\boldsymbol{V} = \boldsymbol{I}$

While as above past research has generally favoured DWLS to MLR (or WLS), among these three a simulation study suggests that we should favour ULS as our default as it is generally more accurate and efficient than DWLS - albeit with the caveat that DWLS may converge in situations where ULS cannot (Forero, Maydeu-Olivares, and Gallardo-Pujol 2009). I therefore utilise ULS in this paper. Similarly to regular CFA, implementing the unit intercept in ordinal CFA is relatively straightforward. We can either set the loadings of the unit-intercept factor to 1 while freeing its variance:

$$x^*_{ij} = \lambda_{jc}\eta_{ic} + 1\eta_{ia} + \epsilon_{ij} \tag{11}$$

Or alternatively we can can constrain its variance to 1 while constraining the loadings to be equal but freely estimating their value:

$$x^*_{ij} = \lambda_{jc}\eta_{ic} + \lambda_a\eta_{ia} + \epsilon_{ij} \tag{12}$$

Continuing with the convention established above, for the remainder of this paper I refer to these models as (11) OCFA1 and (12) OCFA2.

Fully Unbalanced Scales

Past simulation studies suggest that unit-intercept models are robust to unbalanced scales where other acquiescence-correction methods require balanced scales (Savalei and Falk 2014). However, the crucial point made above is that the unit-intercept requires *contradiction* in order to empirically identify the acquiescence component, which is lacking in *fully unbalanced* scales. If for instance we take the BSA left-right scale, it is impossible to try and tell apart those who are agreeing with left-wing statements because they agree with them and those who are agreeing with the same statements because they are acquiescent. *There is no information available to distinguish the two kinds of agreement.* To solve this problem and empirically identify[4] the acquiescence component, I use Watson's idea of introducing further information in the model (1992). Specifically, if a scale which contains statements for which it would be contradictory to agree to all of them, it can be used to identify the acquiescence component in itself and thus also in the fully unbalanced scale. It is unfortunate that a strategy does not exist based on the fully unbalanced scale alone, but it should be clear that it is not possible to identify acquiescence bias in such a scale *without additional information being introduced in some form.* While the same simulations suggest that the unit intercept CFA model is robust to differing levels of acquiescence bias in each indicator (Savalei and Falk 2014), this approach necessarily strengthens the assumption, as it assumes not only the same level of acquiescence for each respondent on one scale, but on all scales in the model.

Methodology and Data

Simulation

With the various configurations of unit-intercept CFA models established, I now turn to the task of testing the performance of these models. Given most applied researchers will be interested in obtaining measurements of the content factors to use in regression models rather than the specific pattern of loadings in the CFA models,

[4] I use this term to distinguish from statistical identification of the model. The model may well be statistically identified (a unique solution exists), but that is no reason to imagine we have successfully captured the acquiescence component

my emphasis is on the differences in predicted factor scores from these models. I begin with a a Monte Carlo (MC) simulation of the performance of the four models. In each simulation, I simulate three common factors, one independent variable which partially determines acquiescence, an acquiescence factor, and six indicators for each of the three common factors. Where the indicators differ is that for the first common factors, all of the indicators run in the same direction. By contrast, the indicators for the second and third common factors represent balanced scales. However, while those for the second common factor are just as acquiescent as those for the first common factor, those for the third common factor are half as acquiescent as those for the first and second common factors. By specifying the simulations in this manner, it becomes possible to test how robust the methods are to violations of the assumption of common acquiescence across all scales included. The R code for the simulation is available in the supplementary material along with other code for this paper.

I therefore run the four correction methods to extract estimations of the first common factor, with the explicit aim of stripping them of acquiescence bias. Of course, by extension the CFA models will also strip these predictions of their unique components. It is this feature of the CFA models that would make them desirable to use even in the case of balanced scales - it enables the researcher to obtain measures of the latent variable of interest stripped of both noise and bias, and also relaxes the assumption that each indicator captures said variable equally. I therefore predict factor scores from each of these models for the first content factor, and show its correlation with the 'true' factor values and the acquiescence values. For comparison, I also simulate results for CFA and OCFA models without any correction (i.e. containing only the 6 indicators with no attempt to correct the bias) and the correlations with the Likert scale that would otherwise be produced from the indicators. Finally, I regress all of the different measures of the first common factor on the aforementioned independent variable that partially determines acquiescence but is independent of the common factors, to assess the extent to which the models correct biased regression results.

Application

Of course, the performance of these models on simulated data while informative is not a complete picture. I therefore also perform correction on a third dataset, taking the BSA and BES differences as a baseline against which to compare. To do this, I utilise data from the fourteenth wave of the British Election Study Internet Panel (BESIP) as cross-sectional data (Fieldhouse et al. 2020). I choose the fourteenth wave first because it contains fully unbalanced left-right and libertarian-authoritarian scales similar to those above, and second because it contains not one but two balanced scales which can be used to identify the acquiescence component. This was the May 2018 wave of BESIP and thus some comparability to the earlier two surveys is lost, but for my purposes here it remains a good choice for the aforementioned reasons. The wording of the left-right statements are:

- **lr1:** Government should redistribute income from the better off to those who are less well off
- **lr2:** Big business takes advantage of ordinary people
- **lr3:** Ordinary working people do not get teir fair share of the nation's wealth
- **lr4:** There is one law for the rich and one for the poor
- **lr5:** Management will always try to get the better of employees if it gets the chance

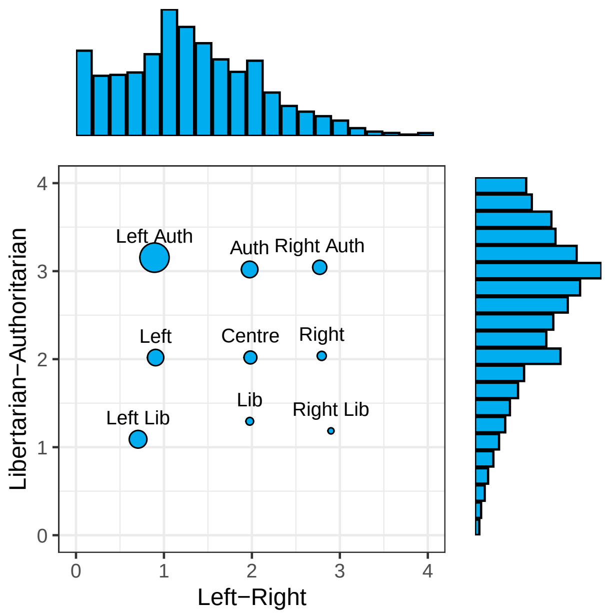
The wordings of the libertarian-authoritarian statements are:

- **al1:** Young people today don't have enough respect for traditional authority
- **al2:** For some crimes, the death penalty is the most appropriate sentence
- **al3:** Schools should teach children to obey authority
- **al4:** Censorship of films and magazines is necessary to uphold moral standards
- **al5:** People who break the law should be given stiffer sentences

As in the BES face to face survey, these questions have five responses ranging from 'strongly disagree' (1) to 'strongly agree' (5). Figure 2 shows the marginal and joint distributions of the 0-4 scales produced from these indicators. Similar to figure 1 above, the x-axis shows the left-right scale and the y-axis shows the libertarian-authoritarian scale. As before opposite these axes are the respective histograms for the scales so as to visualise the marginal distributions. Likewise, to visualise the joint distribution of these scales respondents have been grouped into 9 distinct categories and their means plotted. As a reminder, those with scores ranging from 0 to 1.6 were placed in the 'left' and 'libertarian' groups of the respective dimensions.

Correspondingly, those with scores ranging from 2.4 to 4 were placed in the 'right' and 'authoritarian' groups of the respective dimensions. Finally, those in-between these values were placed in the 'centre' group for each dimension. The location of each group's point corresponds to that groups' means, while the size of the dot corresponds to the number of respondents in that group.

Figure 2: Joint Distribution of BESIP Scales



When compared to 1, it is clear that the distributions of the scales in wave 14 of BESIP are considerably more in line with those in the BSA than those in the BES face to face survey. This is unsurprising - as with the BSA, the BESIP scales are comprised of items that all contain left-wing and authoritarian statements and thus are biased in this direction by acquiescence bias. The marginal and joint distributions are a sufficiently close match to further lend credence to the idea that the earlier BSA-BES differences were in fact driven by acquiescence bias. After performing the simulation, I turn to attempting to correct these scales through the aforementioned unit-intercept CFA methods. This task is split across two balanced scales. These scales cover attitudes to zero-sum approaches to life; and personal empathy. The statements on the zero-sum scale are:

- **zero1:** One person's loss is another person's gain (zero-sum)
- **zero4:** There's only so much to go around. Life is about how big a slice of the pie you can get. (zero-sum)
- **zero5:** Life isn't about winners and losers, everyone can do well (everyone can win)
- **zero7:** The only way to make someone better off is to make someone else worse off (zero-sum)
- **zero9:** There are ways to make everyone better off without anyone losing out (everyone can win)
- **zero11:** Everyone can be a winner at the same time (everyone can win)

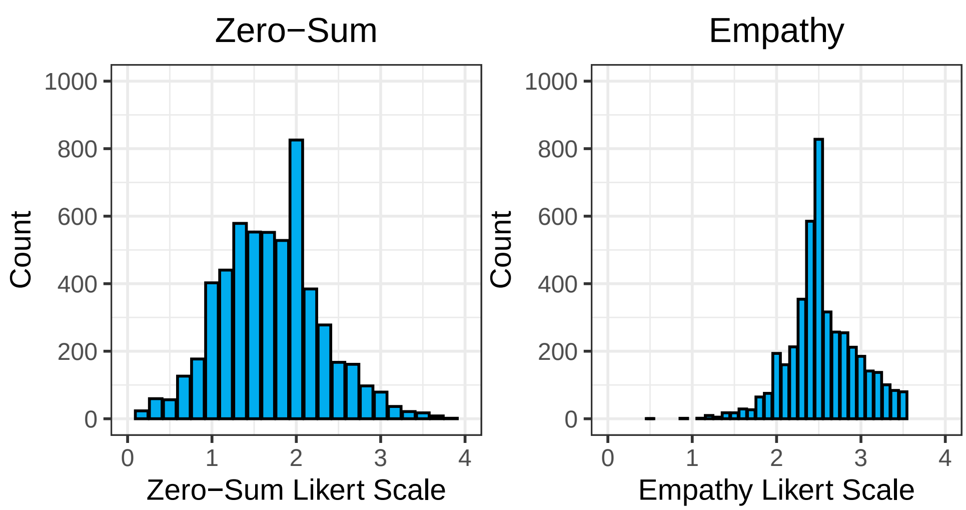
The statements from the empathy scale are:

- **empathy1:** I can usually figure out when my friends are scared (empathetic)
- **empathy2:** I can usually realize quickly when a friend is angry (empathetic)
- **empathy3:** I can usually figure out when people are cheerful (empathetic)
- **empathy4:** I am not usually aware of my friends' feelings (unempathetic)
- **empathy5:** When someone is feeling 'down' I can usually understand how they feel (empathetic)
- **empathy6:** After being with a friend who is sad about something, I usually feel sad (empathetic)
- **empathy7:** My friends' unhappiness doesn't make me feel anything (unempathetic)
- **empathy8:** Other people's feelings don't bother me at all (unempathetic)
- **empathy9:** I don't become sad when I see other people crying (unempathetic)

- **empathy10:** My friends' emotions don't affect me much (unempathetic)

As with the two scales above all of the questions have five responses ranging from 'strongly disagree' (1) to 'strongly agree' (5). One notable feature of the empathy and zero-sum scales in the fourteenth wave is that there is no overlap between them - they were asked in different subsections of the wave. This in essence creates two opportunities to test the performances of the models using different scales to aid the corrections - albeit in different samples. The complete cases (i.e. for both the left-right and libertarian authoritarian scales along with the relevant balanced scale) for these subsamples are 5836 for the zero-sum subset and 4478 for the empathy subset. Figure 3 shows bar plots of the two balanced scales. The zero-sum scale ranges from everyone can win (0) to zero-sum (4), while the empathy scale runs from unempathetic (0) to empathetic (4).

Figure 3: Bar Plots of the Balanced Scales



Although analysis of these scales is not the focal point of this paper, it is worth taking a moment to discuss a few points. First, the empathy scale is notably less dispersed than the other scales discussed in this paper. This may be a function of the fact that it is comprised of a higher number of indicators than any of the others (10, as opposed to 5 or 6). It may also be however that given individuals are generally predisposed to view themselves as empathetic that there is less noise - and overall acquiescence - in the empathy scale. To test this second point, I ran CFA1 models on each of these scales alone (i.e. a correction model for balanced data with no other scale indcluded), the full results for which are available in the appendix in tables **??** and **??**. Of particular relevance are the respective variance components. In the zero-Sum model, the variance of the acquiescence component is 0.099, while in the empathy model the acquiescence component has the smaller value of 0.05. This suggests that there is in fact less acquiescence in the empathy scale than in the zero-sum scale, which could carry implications for attempted corrections.

As in the simulation study, my emphasis is on obtaining corrected measures of these scales as this is likely of most interest to applied researchers. I therefore emphasise predicted factor scores from these models, rather than any particular set of model parameters. In both cases, I run all four correction methods to obtain corrected scores, and provide visualisations of the joint distributions of these scores (along with visualisations of the marginal distributions in the appendix) in order to examine how distributional inferences may change from the Likert scales earlier. From here, I regress the raw likert scales in the subsamples along with the four measurements on an education level variable to examine if and how regression results change with the correction methods. To run the CFA models in this paper, R version 4.1.0 and lavaan version 0.6-9 (Rosseel 2012) with code adapted from the appendix of Savalei et al (2014). Other R packages used throughout this paper for either substantive results or their presentation (including in the demonstration section) are texreg version 1.37.5 (Leifeld 2013), semTable version 1.8, tidyverse version 1.3.1, ggpubr version 0.4.0, gtable version 0.3.0, gridExtra 2.3, ggcorrplot version 0.1.3, survey version 4.0, srvyr version 1.0.1, and jtools version 2.1.4. Since lavaan does not support weights for ordinal CFA models I have not used them in the CFA models themselves, but they were used in producing distributions from the predicted factor scores of the models. Since the associations between variables should be reasonably robust to weighting, this is likely unproblematic.

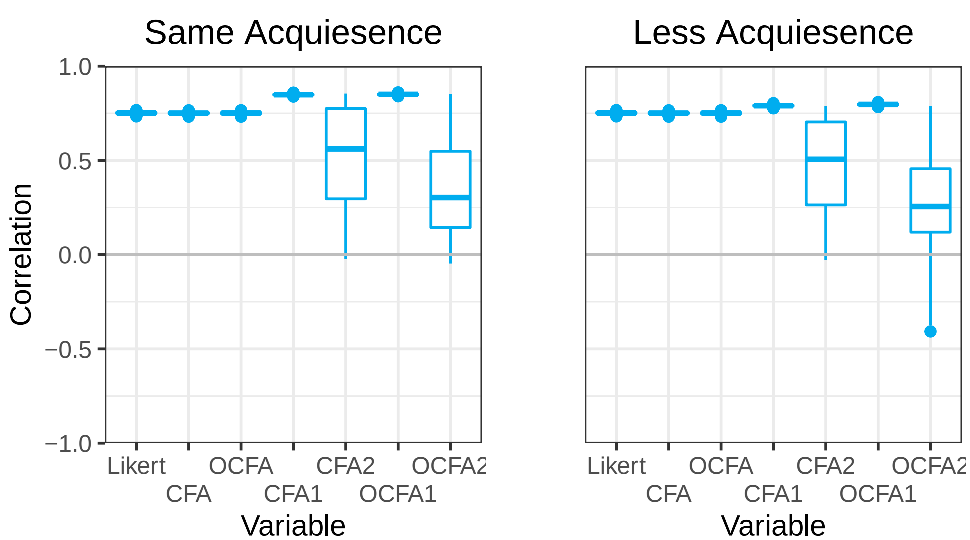
Results

This section is split in two. The first subsection shows the results of the simulation study testing the performance of different versions of the unit-intercept model. The second subsection shows the results of application of these models to the BESIP data. Importantly, this second subsection relies on information gained about the performance of the unit-intercept CFA variants from the simulation study.

Simulation Study

Perhaps the most important feature of any correction method is its ability to obtain reasonably good measurements of the concept of interest. To assess this, figure 4 shows graphs containing boxplots of the correlation coefficients between potential measurement options and the 'true' factor values of the first common factor from the simulation (i.e. of the fully unbalanced scale simulated). The graph on the left contains those from the simulation where the identifying scale contained the same level of acquiescence, while the graph on the right contains those from the simulation where the identifying scale contained half the amount of acquiescence. I retain this column placement across future plots. These plots further include the correlation coefficient between the Likert scale produced from the indicators and the true facot score to provide a baseline against which to compare the performance of the correction models. I also include two CFA models - one unordered estimated with MLR and one ordered estimated with ULS - to show how these models without any attempt at correction might compare. 'CFA' and 'OCFA' without numbers refer to CFA models of the relevant kind without any attempt to correct for acquiescence.

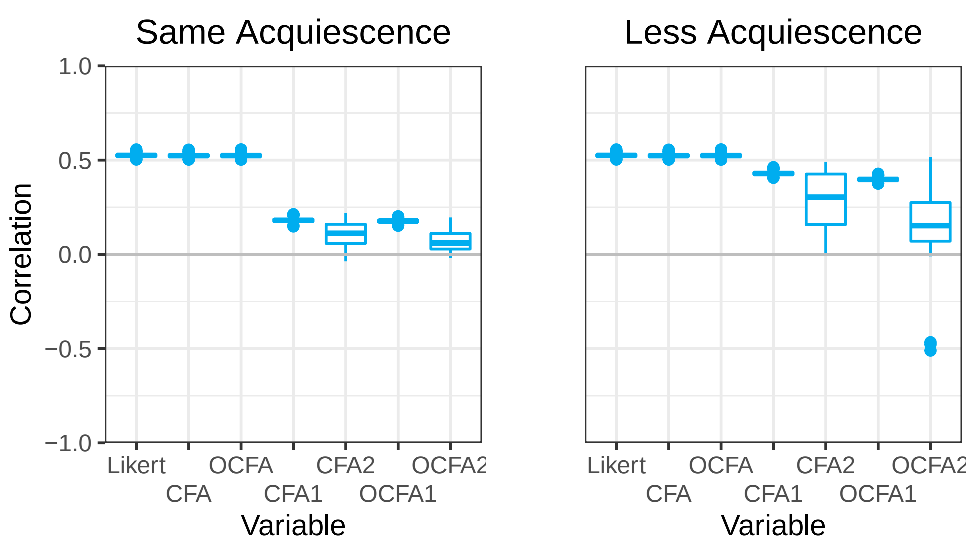
Figure 4: Measure Correlations with True Factor Values



As is visible in figure 4, the CFA1 and OCFA1 models outperform both the Likert scales and CFA and OCFA models without any attempt at correction in terms of successfully measuring the 'true' factor. They consistently have a higher correlation with it, even where the assumption of equal acquiescence across scales is violated. The violation of this assumpt *does* reduce the gap in performance, but it remains the case that even in non-ideal conditions these versions of the unit-intercept perform well. However, the CFA2 and OCFA2 models perform worse by comparison, with a large range for the correlations showing that not only is the central tendency of their correlations *worse* than the Likert scale or CFA/OCFA models with no correction, but they are highly inconsistent estimators. There is not a clear reason why the CFA2 and OCFA2 models should be inconsistent measures. It is however clear that - at least as far as this simulation study is concerned - that treating the indicators as ordinal as opposed to continuous does not appear to make much of a difference. This being said, the simulation does not offer a particularly strenuous test of this. The models perform similarly in terms of their ability to capture the acquiescence factor, which is visualised in figure B1 in the appendix.

It is clear that insofar as these simulation results can show CFA2 and OCFA2 are noisy estimators relative to CFA1 and OCFA1. The second question of interest to substantive researchers however are the ability of these versions of the unit-intercept CFA to produce *unbiased* estimates - i.e. estimates with the acquiescence component scrubbed out. To this end, figure 5 shows the correlations between the various measurement options for the first content factor and the true level of acquiescence. The more successful these correction methods are, the more acquiescence that will be removed and thus the closer to 0 these correlations should be. Once again I include the likert scale constructed from the indicators and CFA and OCFA methods without any correction to serve as a baseline for comparison.

Figure 5: Measure Correlations with True Acquiescence



Some important features emerge from figure 5. First, relative to raw Likert Scales and CFA methods without correction, all four versions of the unit intercept CFA model produce less biased results. However, just as the improved measurement quality does depend on the extent to which the assumption of equal acquiescence across all scales is met, so too does the ability of these models to remove acquiescence bias. The harm caused is much more noticeable - the CFA1 and OCFA1 models shift much closer in their correlations with acquiescence towards those of the Likert scale and the correction-free CFA and OCFA models, while the CFA2 and OCFA2 also have smaller shifts in the same direction and they become more inconsistent in their ability to remove acquiescence bias. However, notably these models do perform better than CFA1 and OCF1 at removing acquiescence bias - OCFA2 in particular performs reasonably well. Taken together, these results suggest that CFA1 and OCFA1 are *consistent* estimators of the true content factor - but CFA2 and OCFA2 while more inconsistent are better at reducing bias in the measurements of the content factor. This carries implications for both the midpoints of measurements and the regression coefficients with variables related to acquiescence. Figure B2 in the appendix visualises the midranges of the measurements, while figure B3 shows the regression coefficients of an independent variable that causes acquiescence but not the true factor of interest. Both sets of results do suggest that in particular where the assumption of equal acquiescence is violated, the extent to which CFA1 and OCFA1 correct estimates of these quantities is reduced.

Application

With the relative performance of the four variations of the unit-intercept confirmatory factor analysis model now established, I turn to applying them to the fourteenth wave of BESIP. In this subsection I present a series of results demonstrating the comparative performance of the methods. Since my emphasis as throughout the paper is on obtaining corrected measurements, the plots presented here pertain to the predicted factor scores. Tables containing full results for the CFA models can be found in section D of the appendix which is dedicated solely to presenting these results.

It is worthwhile to begin by examining the resulting distributions of the measurements. Figure 6 shows two-dimensional binplots of the resultant measurements from the four variations and the two subsets of the BESIP dataset. The extracted measures were rescaled to range from 0 to 4 to facilitate comparability with the earlier distribution of the Likert scale in BESIP in figure \ref{{fig:besipJoint}. Furthermore, the 'left' factor extracted was multiplied by -1 to make it range from left to right. Libertarian-Authoritarian factors are on the y-axis, where 0 is most libertarian and 4 is most authoritarian. Left-Right factors are on the x-axis, with 0 similarly being the most left-wing and 4 being the most right-wing. The the colour of the bins change from light blue to dark blue as the count of respondents in that bin increases. The plots are organised in columns for correction method, and by row for which of the two BESIP subsets the data are from. This column-row placement is retained for plots showing results from the substantive application of these models. The marginal distributions of the predicted factor scores can be viewed in appendix figures C1 and C2 for the left-right and libertarian-authoritarian scales respectively.

Figure 6: Two-Dimensional Bin Plot of Voter Beliefs

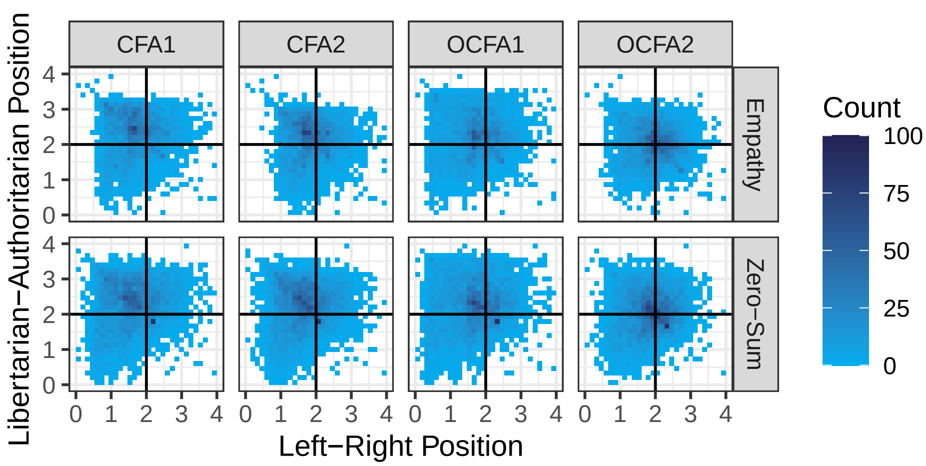
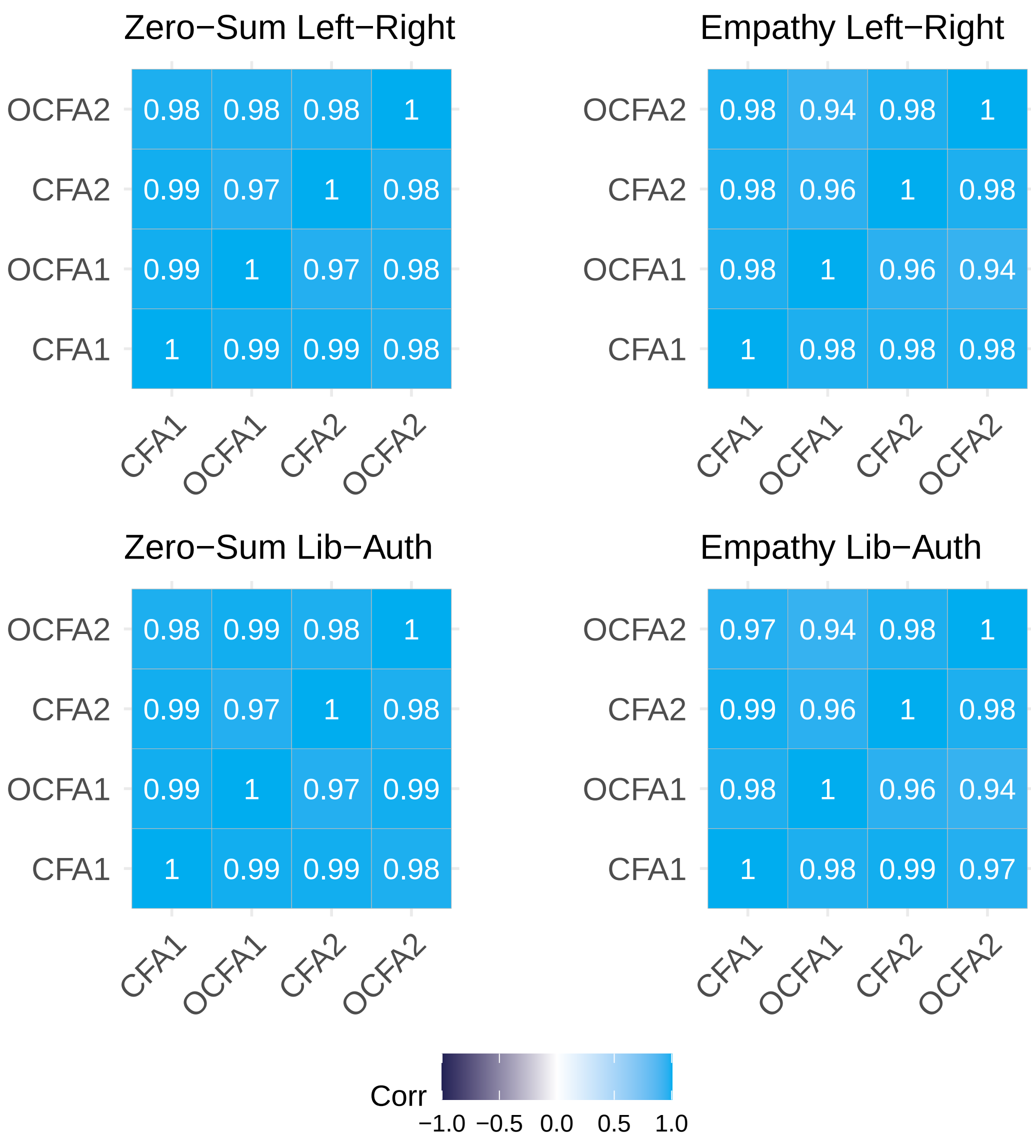


Figure 6 shows that for all of the correction methods, relative to figure 2 there is a shift towards the center of both scales. This shift is particularly pronounced with the OCFA2-obtained measures. Given that the simulation results above suggest that OCFA2 is better at removing acquiescence bias and *all* methods move the scales towards the center, this could imply that in practice for both scales the true positions of survey respondents lie much closer to the center. A table showing the (unweighted) means of the corrected vs likert scales for each subset-model combo is available in appendix table **??** so as to provide a clear numeric confirmation of these differences. Of course, the open question remains as to how well the OCFA2 measure may have captured the content of interest given its performance in the simulation. One way this can be checked is by examining its correlation with the other extracted measures - particularly CFA1 and OCFA1 given their much stronger performances. Figure 7 shows the correlations between the various extracted measures.

Figure 7: Correlations of Predicted Factors



The correlations between the extracted scales are all very near 1, strongly suggesting not only that they are all measuring roughly the same thing - but also that the CFA2 and OCFA2 models may have struggled much more with the simulated data than with real-world data. Given this has been the case for all four scales, it is unlikely to be entirely coincidental - but it is also unclear what aspect of the simulated data if any was particularly difficult for these methods (or alternatively what aspect of the BESIP data that was fine for them). It does however strongly lend credence to the idea that The concentration of respondents around the centers of both scales - as opposed to in the left-authoritarian quadrant - is in line with the true placements of the respondents. This serves as a strong example of how these correction methods could then be used to prevent - or at least reduce - incorrect distributional inferences. These results do of course differ from the BES distribution in figure 1. However, this may be because the models make weaker measurement assumptions than likert scales. Liker scales necessarily implicitly assume either equal acquiescence (in the case of balanced scales) or no acquiescence (in the case of unbalanced scales), equal loadings of the content factor across all items, and equal amounts of noise. By contrast, these CFA methods estimate the acquiescence component, free the loadings of the content factors, and remove all noise. This may be one of the reasons these distributions differ from the BES.

Of course, we are not solely interested in how these corrected measures result in different distributions. We are also interested in the extent towards which they produce corrected regression coefficients. I therefore regress each of these measures, along with the raw likert scales from the samples, on the education level variable available in the BESIP dataset. In these regressions, the measures are always the dependent variable and the reference category for the education level variable is 'no qualifications.' For the results displayed in the main body of the paper, I have avoided recoding the education level variable as this makes some results appear to be somewhat better than they are. I have however included regression results with the recoded education level variable (where the 'below GCSE' category is collapsed into the 'no qualifications' category) in the appendix for more direct comparison. These results are included in appendix tables C4, C5, C6, and C7.

Tables 2 and 3 show the regression results for the left-right scales. The former shows the results for the zero-sum subset, while the latter shows the results for the empathy subset. It is useful to begin here, as this was where the most drastic difference between the BSA and BES results existed. In the shift from the BSA to the BES, the left-right scale shifted from having a positive association with education to a null result. It is against this shift I now compare the following results.

Table 2: Zero-Sum Left-Right

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	1.12***	1.55***	1.61***	1.58***	1.78***
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Below GCSE	0.05	0.04	0.04	0.05	0.04
	(0.07)	(0.05)	(0.05)	(0.06)	(0.05)
GCSE/Equiv	0.09	0.06	0.06	0.08*	0.06
	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)
A-level/Equiv	0.18***	0.10**	0.10**	0.12**	0.09**
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
Undergrad	0.15**	0.05	0.05	0.07	0.03
	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)
Postgrad	0.13*	−0.00	0.00	0.02	−0.04
	(0.06)	(0.04)	(0.04)	(0.05)	(0.04)
R^2	0.00	0.00	0.00	0.00	0.00
Adj. R^2	0.00	0.00	0.00	0.00	0.00
Num. obs.	4965	4965	4965	4965	4965

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

In the case of the results for the zero-sum left-right scales, there is little difference from model to model once the likert scale is dropped. The 'A-level/equiv' category remains significant throughout, and the 'GCSE/equiv' category *becomes* significant in case of the OCFA1 model. However, all models reduce the 'Undergrad' and 'Postgrad' categories to non-significance. It's likely that an applied researcher encountering one of these results would be more likely to conclude in favour of a null interpretation as opposed to rejecting the null - which they plausibly would in the case of the likert scale results. The results for the empathy left-right scales are less differentiated. This is plausibly in part due to the fact that the empathy scale contains less acquiescence per the check mentioned above, but it is also likely due to the fact that as seen in the raw likert scale there appears to be a stronger association with education than in the case of the zero-sum subset. All of the methods to varying degrees reduce the size of the point estimates, but only OCFA2 seems to sufficiently reduce the size of these coefficients such that a researcher would be likelier to conclude in favour of a null result. It is clear that despite its performance in the simulations, OCFA2 does appear to produce results more in line with those in the BES during the demonstration while being highly correlated with the other extracted factors. Why this difference in performances has emerged is not clear.

Table 3: Empathy Left-Right

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	1.08^{***}	1.62^{***}	1.75^{***}	1.59^{***}	1.87^{***}
	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)
Below GCSE	0.17^{*}	0.09	0.09	0.12	0.07
	(0.08)	(0.06)	(0.05)	(0.06)	(0.05)
GCSE/Equiv	0.14^{**}	0.09^{*}	0.10^{**}	0.10^{*}	0.06
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
A-level/Equiv	0.18^{***}	0.11^{**}	0.11^{**}	0.11^{*}	0.07
	(0.05)	(0.04)	(0.04)	(0.05)	(0.04)
Undergrad	0.22^{***}	0.14^{***}	0.15^{***}	0.14^{**}	0.09^{**}
	(0.05)	(0.04)	(0.04)	(0.04)	(0.04)
Postgrad	0.13^{*}	0.08	0.10^{*}	0.06	0.03
	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)
R^2	0.01	0.00	0.00	0.00	0.00
Adj. R^2	0.00	0.00	0.00	0.00	0.00
Num. obs.	3847	3847	3847	3847	3847

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Tables 4 and 5 show the regression results for the left-right scales. As before, the former shows the results for the zero-sum subset, while the latter shows the results for the empathy subset. In the BSA vs BES results above, there was less of a drastic difference in regression results. However, the size of the point estimates for the BES had been noticeably reduced and it is against this baseline which I now compare these results.

Table 4: Zero-Sum Libertarian-Authoritarian

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	3.05^{***}	2.58^{***}	2.51^{***}	2.62^{***}	2.35^{***}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Below GCSE	-0.05	-0.05	-0.05	-0.06	-0.06
	(0.07)	(0.05)	(0.05)	(0.05)	(0.05)
GCSE/Equiv	-0.13^{**}	-0.09^{**}	-0.09^{**}	-0.11^{**}	-0.09^{**}
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
A-level/Equiv	-0.45^{***}	-0.31^{***}	-0.30^{***}	-0.33^{***}	-0.28^{***}
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
Undergrad	-0.76^{***}	-0.53^{***}	-0.52^{***}	-0.57^{***}	-0.47^{***}
	(0.05)	(0.04)	(0.03)	(0.04)	(0.03)
Postgrad	-1.15^{***}	-0.79^{***}	-0.76^{***}	-0.82^{***}	-0.67^{***}
	(0.06)	(0.04)	(0.04)	(0.04)	(0.04)
R^2	0.15	0.13	0.13	0.13	0.12
Adj. R^2	0.15	0.13	0.13	0.13	0.12
Num. obs.	4965	4965	4965	4965	4965

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Given these two results are fairly similar in the differences between the models using predicted factor scores as opposed to the models run with likert scales, they can be discussed together. In both cases, all of the models result in fairly substantial reductions in the absolute size of the point estimates while retaining both significance and the same overall direction. Many of the coefficients in the corrected models - and particularly that of OCFA2 - are similar in size to those in the BES results earlier in table 1. It is clear then that these methods not only perform reasonably well in correcting type one errors, but also appropriately reduce the size of point estimates without causing type two errors. Taken together, this shows the that the unit intercept CFA variations are powerful tools in correcting biases in research results. There remains some ambiguity in these results given the contrast between the performance of these models in the simulation in in this

Table 5: **Empathy Libertarian-Authoritarian**

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	3.11^{***}	2.60^{***}	2.44^{***}	2.59^{***}	2.29^{***}
	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)
Below GCSE	-0.07	-0.03	-0.04	-0.06	-0.02
	(0.08)	(0.05)	(0.05)	(0.06)	(0.05)
GCSE/Equiv	-0.14^{*}	-0.09^{*}	-0.10^{**}	-0.11^{**}	-0.07^{*}
	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)
A-level/Equiv	-0.53^{***}	-0.34^{***}	-0.34^{***}	-0.37^{***}	-0.27^{***}
	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)
Undergrad	-0.77^{***}	-0.51^{***}	-0.50^{***}	-0.55^{***}	-0.41^{***}
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
Postgrad	-1.24^{***}	-0.85^{***}	-0.82^{***}	-0.88^{***}	-0.67^{***}
	(0.06)	(0.05)	(0.04)	(0.05)	(0.04)
R^2	0.16	0.16	0.16	0.15	0.14
Adj. R^2	0.16	0.16	0.16	0.15	0.14
Num. obs.	3847	3847	3847	3847	3847

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

section - particularly for the OCFA2 variant. I now turn to discussing some substantive recommendations and conclusions with this ambiguity in mind.

Conclusion and Reccomendations

It has always been the case that applied researchers need to be alive to *all* parts of the data generating process for their data when performing research. What I hope the results in this paper will have demonstrated is that this is not only in terms of the variables we are substantively interested in, but also in terms of the measurement process for these variables. Public opinion researchers in particular must be alive to potential sources of bias in the surveys we use, especially given the absence of 'objective' measures of our variables of interest. Acquiescence bias in particular is a considerable threat when using likert scales to measure voter attitudes.

Where reseachers are using 'off the shelf' survey data created by others, they should first examine whether the scales are balanced or unbalanced. While I have focused on acquiescent scales in UK data due to the fact that similar scales are reasonably common across several datasets, this problem is by no means constrained to the UK. Some examples of datasets containing unbalanced - partially or fully - likert scales (or indeed agree-disagree items not part of any scale) include but by no means are limited to the Comparative Study of Electoral Systems, the American National Election Survey (in part through the CSES), the European Social Survey, and the World Values survey - to take some prominent examples of popularly used datasets.

When researchers use these likert scales without due consideration of the role of acquiescence bias, they implicity make strong measurement assumptions. Where scales are unbalanced and no word is given to how acquiescence may have biased the results, they are de factor assuming no acquiescence in said results. Moreover, when scales are both balanced and unbalanced, they are assuming all indicators equally capture the content factor in which they are interested, and thus all also contain unique components of equal size. In practice some indicators will better capture the content factor than others.

The unit-intercept CFA methods in their variations allow for these measurement assumptions to be relaxed or outright dropped. While they conceptualise acquiescence as essentially a form of differential person functioning (DPF) constant across all indicators, they freely estimate the loadings of the content factors on these indicators and thus also allow for the unique components to be freely estimated. They are reasonably robust to violations of their assumption of equal acquiescence across all scales. My contribution has been to further clarify the different versions of unit intercept CFA and to develop a strategy for identifying such a model in the case of fully unbalanced likert scales, which one their own cannot be corrected. There is no clear case where the use of these models if possible is not preferable to a raw likert scale.

However, some imperfections and ambiguities remain in these methods. First and most obviously, if correction is being applied to the case of fully unbalanced data, a partially unbalanced or balanced likert scale must be available to identify the acquiescenece bias within. There is no way of avoiding this kind of limitation - without some external information, it will be impossible to know which respondents have agreed because they agree with an item and which are agreeing through acquiescence. It may be the case that other identification approaches are possible (other indicators of the same factor measured through a different method may be a promising route), but the fundamental limitation remains. While given the purpose of my paper the simulations have focussed only on fully unbalanced scales, they do raise questions about the relative performance of different specifications of the unit intercept CFA model. The simulations strongly favour the traditional version of the model, but the applied research results in this paper suggest that something may have gone wrong in the simulation for this model that does not go wrong with real-world data.

Applied researchers thus need to balanced several considerations. First, the degree to which obtaining unbiased measures is central to their reserach. Where these scales are a central variable - either as a depedent variable or a key explanatory variable - it is clear that if possible correction methods should be applied. At best, in the case of balanced scales where the assumptions are likely correct acquiescence bias will serve as a form of additional noise. Where these scales are a key independent variable, this will result in attentuation bias on the parameter estimates. At worst, acquiescence will serve as essentially a known case of omitted variable bias. If these correction methods cannot be applied, researchers should at minimum explicitly discuss how acquiescence bias may have resulted in biases in their results.

Because of the ambiguities and imperfections in these correction methods however, it remains the case that as much as possible acquiescence bias should be dealt with at the survey design stage. Given I am making this recommendation, it is important to consider some potential objections that may arise. Some survey designers in the UK may have used unbalanced scales because past comparisons for the left-right and libertarian-authoritarian scales explored in this paper suggested that they had greater reliability relative to balanced scales (see G. A. Evans and Heath 1995). This is true - just as it is true that as the same research notes they are orthogonal where balanced scales were correlated, that the unbalanced scales had greater associations with various demographic variables. As the authors note, *all* of these differences are due to acquiescence bias.

A more serious challenge emerges in a subset of the literature which suggests fully unbalanced scales *should* be used (see e.g. Swain, Weathers, and Niedrich 2008). In short, the point this article and others makes is that misresponse is most likely due to cognitively difficult tasks from reverse-worded items - e.g. inclusion of the word 'not' to negate a given item. I do not dispute their results - the problems in simply negating item wordings are well-established. I do dispute however the notion that fully unbalanced scales are a good solution - the results in this paper should hopefully put rest to the notion that this is the case. My reccomendation to survey designers is therefore to use *polar opposite* statements, rather than simple *negations*. In particular, this also aids an additional reccomendation - where survey designers are working on long-term studies (such as the BSA, BES, and BESIP), it may well be preferable to simply insert additional items (in these cases three will be enough) to balance out future likert scales while continuing to allow for cross-temporal comparison with historical data. This will not only be of use to substantive applied research, but may also be helpful in future research further exploring acquiescence bias - and its correction.

Bibliography

Billiet, Jaak B, and Eldad Davidov. 2008. "Testing the Stability of an Acquiescence Style Factor Behind Two Interrelated Substantive Variables in a Panel Design." *Sociological Methods & Research* 36 (4): 542–62.

Billiet, Jaak B, and McKee J McClendon. 2000. "Modeling Acquiescence in Measurement Models for Two Balanced Sets of Items." *Structural Equation Modeling* 7 (4): 608–28.

Brown, Timothy A. 2015. *Confirmatory Factor Analysis for Applied Research.* Guilford publications.

Cloud, Jonathan, and Graham M Vaughan. 1970. "Using Balanced Scales to Control Acquiescence." *Sociometry*, 193–202.

Evans, Geoffrey A, and Anthony F Heath. 1995. "The Measurement of Left-Right and Libertarian-Authoritarian Values: A Comparison of Balanced and Unbalanced Scales." *Quality and Quantity* 29 (2): 191–206.

Evans, Geoffrey, Anthony Heath, and Mansur Lalljee. 1996. "Measuring Left-Right and Libertarian-Authoritarian Values in the British Electorate." *British Journal of Sociology*, 93–112.

Fieldhouse, E., J. Green, G. Evans, J. Mellon, and C. Prosser. 2020. "British Election Study Internet Panel." https://www.britishelectionstudy.com/data-objects/panel-study-data/.

Fieldhouse, E., J. Green, G. Evans, H. Schmitt, C. van der Eijk, J. Mellon, and C. Prosser. 2017. "British Election Study 2017 Face-to-Face Post-Election Survey." https://www.britishelectionstudy.com/data-objects/cross-sectional-data/.

Forero, Carlos G, Alberto Maydeu-Olivares, and David Gallardo-Pujol. 2009. "Factor Analysis with Ordinal Indicators: A Monte Carlo Study Comparing DWLS and ULS Estimation." *Structural Equation Modeling* 16 (4): 625–41.

Heath, Anthony, Geoffrey Evans, and Jean Martin. 1994. "The Measurement of Core Beliefs and Values: The Development of Balanced Socialist/Laissez Faire and Libertarian/Authoritarian Scales." *British Journal of Political Science* 24 (1): 115–32.

Kenny, David A, and Deborah A Kashy. 1992. "Analysis of the Multitrait-Multimethod Matrix by Confirmatory Factor Analysis." *Psychological Bulletin* 112 (1): 165.

Leifeld, Philip. 2013. "Texreg: Conversion of Statistical Model Output in r to LATEX and HTML Tables." *Journal of Statistical Software* 55 (8): 1–24.

Li, Cheng-Hsien. 2016. "Confirmatory Factor Analysis with Ordinal Data: Comparing Robust Maximum Likelihood and Diagonally Weighted Least Squares." *Behavior Research Methods* 48 (3): 936–49.

Maydeu-Olivares, Albert, and Donna L Coffman. 2006. "Random Intercept Item Factor Analysis." *Psychological Methods* 11 (4): 344.

Mirowsky, John, and Catherine E Ross. 1991. "Eliminating Defense and Agreement Bias from Measures of the Sense of Control: A 2 x 2 Index." *Social Psychology Quarterly*, 127–45.

NatCen-Social-Research. 2017. "British Social Attitudes Survey, 2017." https://www.ukdataservice.ac.uk/.

Primi, Ricardo, Nelson Hauck-Filho, Felipe Valentini, Daniel Santos, and Carl F Falk. 2019. "Controlling Acquiescence Bias with Multidimensional IRT Modeling." In *The Annual Meeting of the Psychometric Society*, 39–52. Springer.

Primi, Ricardo, Daniel Santos, Filip De Fruyt, and Oliver P John. 2019. "Comparison of Classical and Modern Methods for Measuring and Correcting for Acquiescence." *British Journal of Mathematical and Statistical Psychology* 72 (3): 447–65.

Ray, John J. 1979. "Is the Acquiescent Response Style Problem Not so Mythical After All? Some Results from a Successful Balanced f Scale." *Journal of Personality Assessment* 43 (6): 638–43.

Rosseel, Yves. 2012. "Lavaan: An r Package for Structural Equation Modeling and More. Version 0.5–12 (BETA)." *Journal of Statistical Software* 48 (2): 1–36.

Savalei, Victoria, and Carl F Falk. 2014. "Recovering Substantive Factor Loadings in the Presence of Acquiescence Bias: A Comparison of Three Approaches." *Multivariate Behavioral Research* 49 (5): 407–24.

Surridge, Paula. 2018. "The Fragmentation of the Electoral Left Since 2010." *Renewal: A Journal of Labour Politics* 26 (4): 69–78.

Swain, Scott D, Danny Weathers, and Ronald W Niedrich. 2008. "Assessing Three Sources of Misresponse to Reversed Likert Items." *Journal of Marketing Research* 45 (1): 116–31.

Ware Jr, John E. 1978. "Effects of Acquiescent Response Set on Patient Satisfaction Ratings." *Medical Care*, 327–36.

Watson, Dorothy. 1992. "Correcting for Acquiescent Response Bias in the Absence of a Balanced Scale: An Application to Class Consciousness." *Sociological Methods & Research* 21 (1): 52–88.

Winkler, John D, David E Kanouse, and John E Ware. 1982. "Controlling for Acquiescence Response Set in Scale Development." *Journal of Applied Psychology* 67 (5): 555–61.

Yang-Wallentin, Fan, Karl G Jöreskog, and Hao Luo. 2010. "Confirmatory Factor Analysis of Ordinal Variables with Misspecified Models." *Structural Equation Modeling* 17 (3): 392–423.

Appendix A - Demonstration

Table A1: BSA Education Recode

Original Coding	New Coding
Postgraduate degree	Postgrad
First degree	Undergrad
Higher educ below degree	A-level/equiv
A level or equiv	A-level/equiv
O level or equiv	GCSE/equiv
CSE or equiv	GCSE/equiv
Foreign or other	Missing
No qualification	No Qualification

Table A2: BES Education Recode

Original Coding	New Coding
No qualifications	No qualification
Below GCSE	No qualification
GCSE	GCSE/equiv
A-level	A-level/equiv
Undergraduate	Undergrad
Postgrad	Postgrad

Indicators common to both datasets:

- **Ind1:** There is one law for the rich and one for the poor
- **Ind2:** Young people today don't have enough respect for traditional British values
- **Ind3:** Censorship of films and magazines is necessary to uphold moral standards
- **Ind4:** For some crimes, the death penalty is the most appropriate sentence
- **Ind5:** People who break the law should be given stiffer sentences

Reference category: July

Appendix B - Simulation

Because all CFA models assume the mean of the latent variables to be 0 for purposes of identification, it is not informative to examine their means. Instead, the entire distributions will be shifted *around* this mean. If there is a bias in the positive direction, the midranges (i.e. literally the midpoint of the scale) will therefore be biased in a negative direction. Figure B2 therefore shows the midranges of the predicted factor scores from all six models. It further includes the midranges of the true factor to offer a comparison.

Table A3: Regression of Survey Membership on Common Indicators

	OLS	Logit	Probit
Intercept	0.56^{***}	0.25^{*}	0.16^{*}
	(0.03)	(0.11)	(0.07)
Ind1	0.01	0.05	0.03
	(0.01)	(0.03)	(0.02)
Ind2	−0.00	−0.02	−0.01
	(0.01)	(0.03)	(0.02)
Ind3	−0.01	−0.05	−0.03
	(0.01)	(0.03)	(0.02)
Ind4	-0.02^{**}	-0.06^{**}	-0.04^{**}
	(0.01)	(0.02)	(0.01)
Ind5	0.04^{***}	0.16^{***}	0.10^{***}
	(0.01)	(0.04)	(0.02)
R^2	0.01		
Adj. R^2	0.00		
Num. obs.	5170	5170	5170
AIC		7039.76	7040.06
BIC		7079.07	7079.37
Log Likelihood		−3513.88	−3514.03
Deviance		6841.30	6841.57

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table A4: Regression of Scales on Survey Month

	Left-Right	Lib-Auth
Intercept	1.64^{***}	2.04^{***}
	(0.02)	(0.02)
Aug	−0.03	0.00
	(0.03)	(0.03)
Sep	−0.04	0.04
	(0.04)	(0.05)
R^2	0.00	0.00
Adj. R^2	−0.00	−0.00
Num. obs.	1789	1914

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Appendix C - Application

Appendix D - CFA Results

Table C1: Zero-Sum CFA Check

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
Zero				
zero1	0.41	0.01	27.93	.000
zero4	0.49	0.02	30.94	.000
zero5	-0.53	0.02	-34.32	.000
zero7	0.61	0.01	45.32	.000
zero9	-0.59	0.01	-42.79	.000
zero11	-0.58	0.02	-37.75	.000
Acq				
zero1	1.00^{+}			
zero4	1.00^{+}			
zero5	1.00^{+}			
zero7	1.00^{+}			
zero9	1.00^{+}			
zero11	1.00^{+}			
		Residual Variances		
zero1	0.55	0.01	40.40	.000
zero4	0.66	0.02	38.07	.000
zero5	0.55	0.02	30.70	.000
zero7	0.38	0.02	24.65	.000
zero9	0.45	0.02	28.09	.000
zero11	0.62	0.02	35.43	.000
		Latent Variances		
Zero	1.00^{+}			
Acq	0.10	0.00	20.68	.000
		Fit Indices		
χ^2(df)	253.63			
CFI	0.96			
TLI	0.93			
RMSEA	0.07			
Scaled χ^2(df)	181.16(8)			.000

$^{+}$Fixed parameter

Table C2: Empathy CFA Check

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
Empathy				
em1	0.30	0.01	27.71	.000
em2	0.32	0.01	32.32	.000
em3	0.30	0.01	30.87	.000
em4	-0.34	0.01	-25.90	.000
em5	0.29	0.01	28.72	.000
em6	0.25	0.01	21.55	.000
em7	-0.45	0.01	-43.42	.000
em8	-0.48	0.01	-46.42	.000
em9	-0.39	0.01	-29.51	.000
em10	-0.47	0.01	-43.68	.000
Acq				
em1	1.00^{+}			
em2	1.00^{+}			
em3	1.00^{+}			
em4	1.00^{+}			
em5	1.00^{+}			
em6	1.00^{+}			
em7	1.00^{+}			
em8	1.00^{+}			
em9	1.00^{+}			
em10	1.00^{+}			
		Residual Variances		
em1	0.16	0.01	22.44	.000
em2	0.09	0.01	17.78	.000
em3	0.10	0.01	19.15	.000
em4	0.44	0.01	30.44	.000
em5	0.22	0.01	25.58	.000
em6	0.37	0.01	35.33	.000
em7	0.15	0.01	19.24	.000
em8	0.14	0.01	20.37	.000
em9	0.34	0.01	29.49	.000
em10	0.15	0.01	21.08	.000
		Latent Variances		
Empathy	1.00^{+}			
Acq	0.05	0.00	18.78	.000
		Fit Indices		
χ^2(df)	2268.03			
CFI	0.87			
TLI	0.83			
RMSEA	0.12			
Scaled χ^2(df)	1513.06(34)			.000

$^{+}$Fixed parameter

Figure B1: Acquiescence Factor Correlations with True Acquiescence

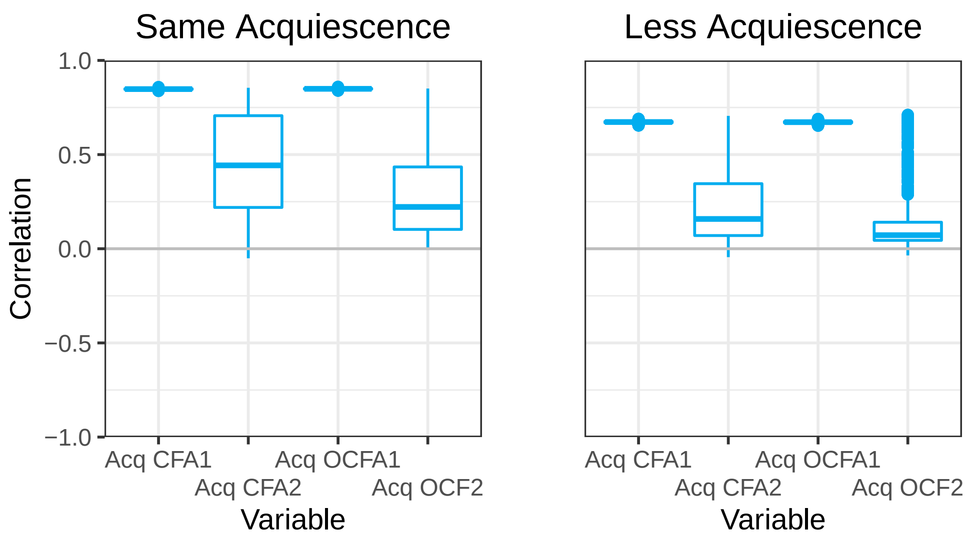


Figure B2: Measure Midranges

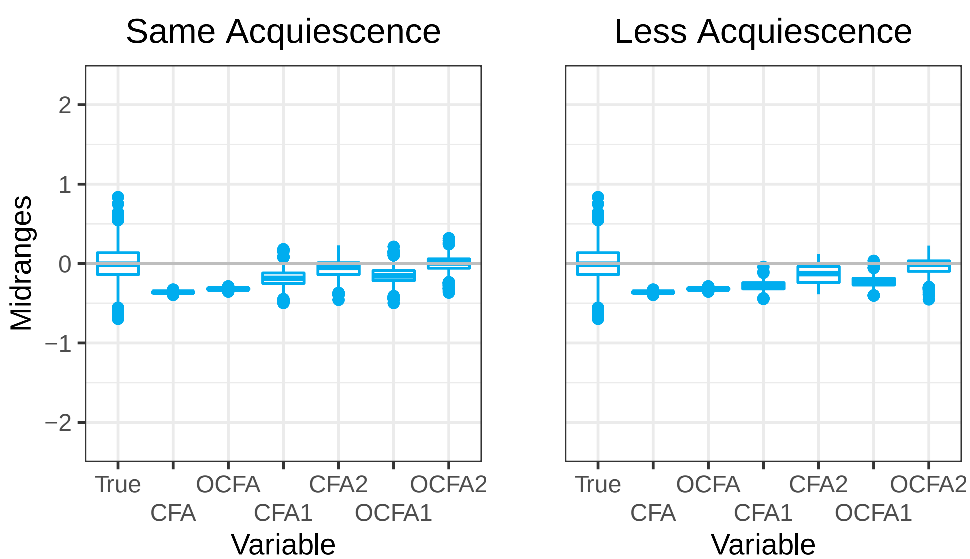


Figure B3: Coefficients from Regression of Measures on Independent Variable

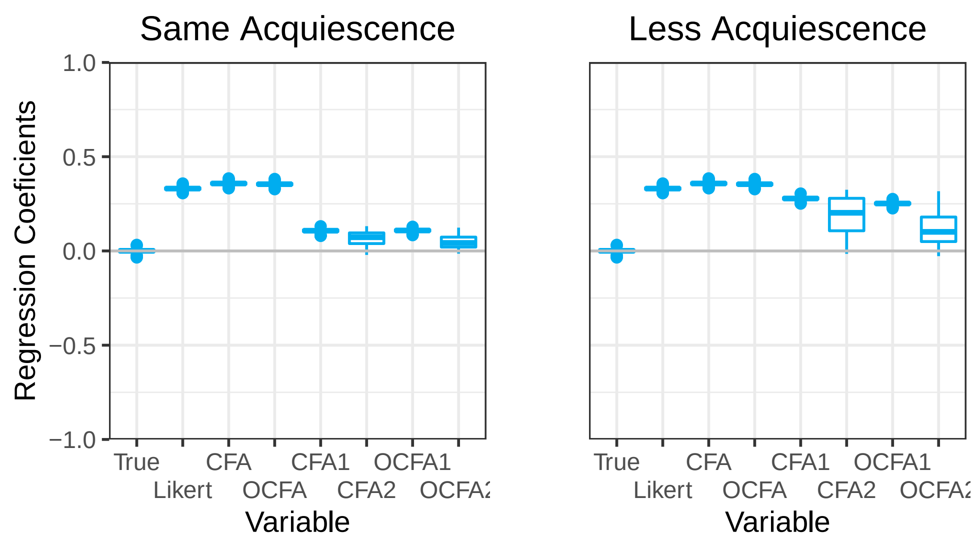


Figure C1: Density Plots of Left-Right Factors from Correction Models

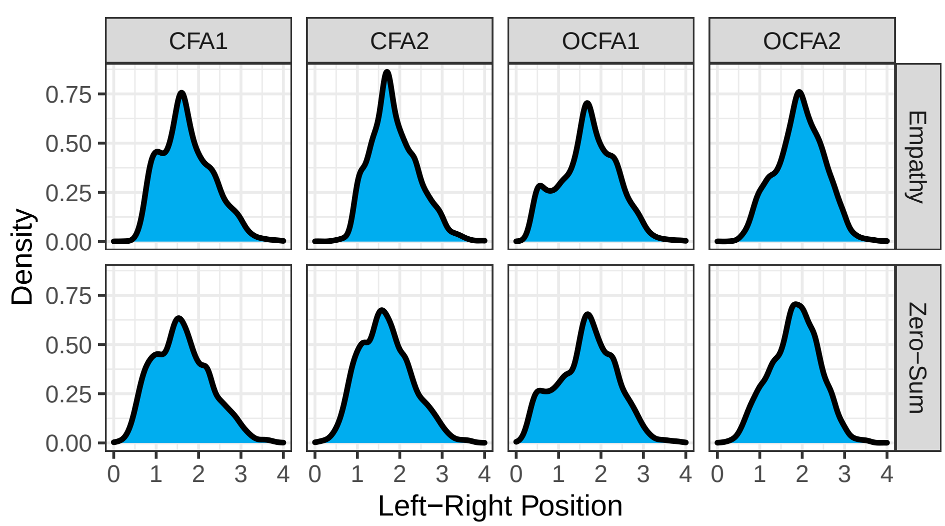


Figure C2: Density Plots of Lib-Auth Factors from Correction Models

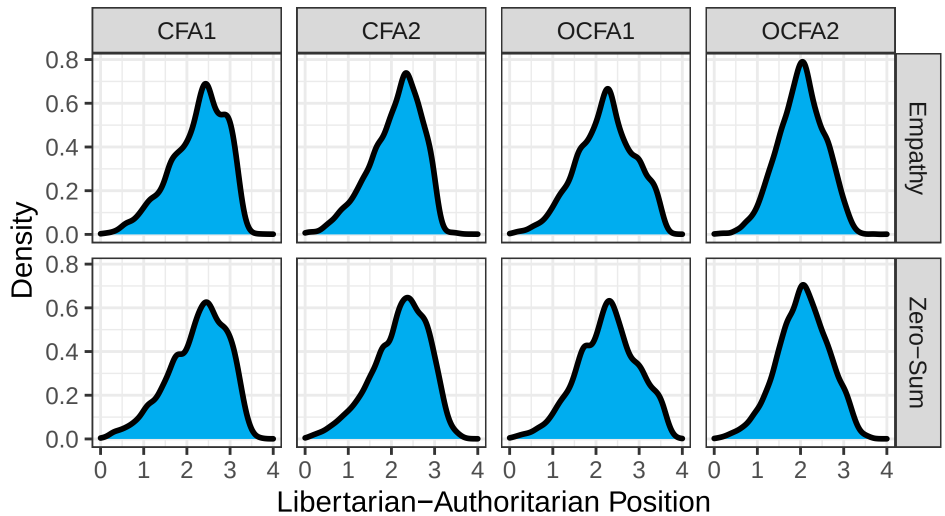


Table C3: Likert vs Factor Means

Subset	Model	Likert Mean LR	Factor Mean LR	Likert Mean LA	Factor Mean LA
Empathy	CFA1	1.25	1.73	2.58	2.26
Empathy	CFA2	1.25	1.87	2.58	2.10
Empathy	OCFA1	1.25	1.69	2.58	2.22
Empathy	OCFA2	1.25	1.94	2.58	2.01
Zero-Sum	CFA1	1.26	1.61	2.57	2.25
Zero-Sum	CFA2	1.26	1.67	2.57	2.19
Zero-Sum	OCFA1	1.26	1.67	2.57	2.26
Zero-Sum	OCFA2	1.26	1.83	2.57	2.06

Table C4: Zero-Sum Left-Right Alternative

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	1.14^{***}	1.56^{***}	1.62^{***}	1.60^{***}	1.80^{***}
	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
GCSE/Equiv	0.07	0.05	0.04	0.06	0.04
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
A-level/Equiv	0.16^{***}	0.09^{**}	0.08^{**}	0.10^{**}	0.07^{*}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Undergrad	0.13^{***}	0.03	0.03	0.05	0.01
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Postgrad	0.12^{*}	-0.02	-0.01	0.00	-0.05
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
R^2	0.00	0.00	0.00	0.00	0.00
Adj. R^2	0.00	0.00	0.00	0.00	0.00
Num. obs.	4965	4965	4965	4965	4965

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table C5: Empathy Left-Right Alternative

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	1.14^{***}	1.65^{***}	1.79^{***}	1.63^{***}	1.90^{***}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
GCSE/Equiv	0.08	0.06	0.06^{*}	0.05	0.04
	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)
A-level/Equiv	0.12^{*}	0.07^{*}	0.08^{*}	0.06	0.04
	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)
Undergrad	0.16^{***}	0.10^{**}	0.11^{***}	0.09^{**}	0.07^{*}
	(0.04)	(0.03)	(0.03)	(0.04)	(0.03)
Postgrad	0.07	0.04	0.06	0.01	0.00
	(0.06)	(0.04)	(0.04)	(0.05)	(0.04)
R^2	0.00	0.00	0.00	0.00	0.00
Adj. R^2	0.00	0.00	0.00	0.00	0.00
Num. obs.	3847	3847	3847	3847	3847

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table C6: Zero-Sum Libertarian-Authoritarian Alternative

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	3.03^{***}	2.56^{***}	2.50^{***}	2.60^{***}	2.33^{***}
	(0.03)	(0.03)	(0.02)	(0.03)	(0.02)
GCSE/Equiv	-0.11^{**}	-0.08^{*}	-0.07^{*}	-0.09^{**}	-0.07^{*}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
A-level/Equiv	-0.44^{***}	-0.29^{***}	-0.28^{***}	-0.31^{***}	-0.26^{***}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Undergrad	-0.74^{***}	-0.52^{***}	-0.50^{***}	-0.55^{***}	-0.45^{***}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Postgrad	-1.13^{***}	-0.77^{***}	-0.74^{***}	-0.80^{***}	-0.65^{***}
	(0.05)	(0.04)	(0.04)	(0.04)	(0.03)
R^2	0.15	0.13	0.13	0.13	0.12
Adj. R^2	0.15	0.13	0.13	0.13	0.12
Num. obs.	4965	4965	4965	4965	4965

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table C7: Empathy Libertarian-Authoritarian Alternative

	Raw	CFA1	CFA2	OCFA1	OCFA2
Intercept	3.08^{***}	2.59^{***}	2.42^{***}	2.57^{***}	2.28^{***}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.02)
GCSE/Equiv	-0.12^{*}	-0.08^{*}	-0.08^{**}	-0.08^{*}	-0.06^{*}
	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)
A-level/Equiv	-0.50^{***}	-0.33^{***}	-0.32^{***}	-0.35^{***}	-0.26^{***}
	(0.05)	(0.03)	(0.03)	(0.04)	(0.03)
Undergrad	-0.74^{***}	-0.50^{***}	-0.48^{***}	-0.53^{***}	-0.40^{***}
	(0.04)	(0.03)	(0.03)	(0.03)	(0.03)
Postgrad	-1.21^{***}	-0.84^{***}	-0.80^{***}	-0.86^{***}	-0.66^{***}
	(0.06)	(0.04)	(0.04)	(0.04)	(0.03)
R^2	0.16	0.16	0.16	0.15	0.14
Adj. R^2	0.16	0.16	0.16	0.15	0.14
Num. obs.	3847	3847	3847	3847	3847

$^{***}p < 0.001$; $^{**}p < 0.01$; $^{*}p < 0.05$

Table D1: Zero-Sum CFA1

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
Z				
zero7	0.58	0.01	41.91	.000
zero1	0.40	0.01	26.74	.000
zero4	0.48	0.02	30.80	.000
zero11	-0.59	0.02	-37.05	.000
zero5	-0.55	0.02	-33.82	.000
zero9	-0.60	0.01	-44.22	.000
LeftCorrected				
lr1	0.81	0.02	52.81	.000
lr2	0.70	0.01	54.76	.000
lr3	0.81	0.01	65.45	.000
lr4	0.83	0.01	65.12	.000
lr5	0.61	0.01	41.30	.000
AuthCorrected				
al1	0.85	0.01	59.46	.000
al2	0.99	0.02	57.09	.000
al3	0.73	0.01	51.48	.000
al4	0.56	0.02	33.15	.000
al5	0.72	0.01	54.43	.000
Acq				
zero7	1.00^{+}			
zero1	1.00^{+}			
zero4	1.00^{+}			
zero11	1.00^{+}			
zero5	1.00^{+}			
zero9	1.00^{+}			
lr1	1.00^{+}			
lr2	1.00^{+}			
lr3	1.00^{+}			
lr4	1.00^{+}			
lr5	1.00^{+}			
al1	1.00^{+}			
al2	1.00^{+}			
al3	1.00^{+}			
al4	1.00^{+}			
al5	1.00^{+}			
		Residual Variances		
zero7	0.42	0.02	27.71	.000
zero1	0.57	0.01	41.82	.000
zero4	0.66	0.02	39.15	.000
zero11	0.64	0.02	35.91	.000
zero5	0.57	0.02	31.08	.000
zero9	0.44	0.02	28.90	.000
lr1	0.77	0.02	37.58	.000
lr2	0.35	0.01	26.28	.000
lr3	0.36	0.02	21.79	.000
lr4	0.45	0.02	28.81	.000
lr5	0.56	0.02	31.96	.000
al1	0.47	0.02	28.16	.000
al2	1.26	0.03	40.44	.000
al3	0.47	0.01	32.28	.000
al4	1.04	0.02	50.58	.000
al5	0.41	0.01	28.89	.000
		Latent Variances		
Z	1.00^{+}			
LeftCorrected	1.00^{+}			

Table D2: Zero-Sum CFA2

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
Z				
zero7	0.67	0.05	12.74	.000
zero1	0.50	0.05	9.47	.000
zero4	0.60	0.05	11.65	.000
zero11	-0.49	0.05	-9.43	.000
zero5	-0.44	0.05	-8.75	.000
zero9	-0.50	0.05	-9.44	.000
LeftCorrected				
lr1	0.83	0.02	48.48	.000
lr2	0.74	0.02	42.63	.000
lr3	0.85	0.01	65.56	.000
lr4	0.87	0.02	49.71	.000
lr5	0.65	0.02	29.13	.000
AuthCorrected				
al1	0.87	0.02	38.85	.000
al2	1.02	0.03	36.27	.000
al3	0.75	0.02	33.41	.000
al4	0.57	0.02	25.19	.000
al5	0.74	0.02	34.50	.000
Acq				
zero7	0.32	0.02	19.80	.000
zero1	0.32	0.02	19.80	.000
zero4	0.32	0.02	19.80	.000
zero11	0.32	0.02	19.80	.000
zero5	0.32	0.02	19.80	.000
zero9	0.32	0.02	19.80	.000
lr1	0.32	0.02	19.80	.000
lr2	0.32	0.02	19.80	.000
lr3	0.32	0.02	19.80	.000
lr4	0.32	0.02	19.80	.000
lr5	0.32	0.02	19.80	.000
al1	0.32	0.02	19.80	.000
al2	0.32	0.02	19.80	.000
al3	0.32	0.02	19.80	.000
al4	0.32	0.02	19.80	.000
al5	0.32	0.02	19.80	.000
		Residual Variances		
zero7	0.43	0.02	26.87	.000
zero1	0.55	0.01	40.40	.000
zero4	0.63	0.02	36.35	.000
zero11	0.63	0.02	35.41	.000
zero5	0.55	0.02	30.24	.000
zero9	0.45	0.02	27.00	.000
lr1	0.77	0.02	34.91	.000
lr2	0.34	0.01	25.38	.000
lr3	0.36	0.02	20.90	.000
lr4	0.45	0.02	28.41	.000
lr5	0.56	0.02	31.64	.000
al1	0.47	0.02	28.77	.000
al2	1.24	0.03	40.28	.000
al3	0.46	0.01	32.63	.000
al4	1.04	0.02	50.89	.000
al5	0.41	0.01	29.59	.000
		Latent Variances		
Z	1.00^{+}			
LeftCorrected	1.00^{+}			

Table D3: Zero-Sum OCFA1

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
Z				
zero7	0.70	0.01	80.09	.000
zero1	0.45	0.01	43.49	.000
zero4	0.52	0.01	50.62	.000
zero11	-0.57	0.01	-62.12	.000
zero5	-0.59	0.01	-62.59	.000
zero9	-0.67	0.01	-79.09	.000
LeftCorrected				
lr1	0.67	0.01	87.21	.000
lr2	0.81	0.01	144.90	.000
lr3	0.83	0.01	154.79	.000
lr4	0.81	0.01	145.17	.000
lr5	0.67	0.01	87.98	.000
AuthCorrected				
al1	0.80	0.01	118.71	.000
al2	0.70	0.01	83.49	.000
al3	0.75	0.01	101.38	.000
al4	0.50	0.01	45.76	.000
al5	0.79	0.01	110.86	.000
Acq				
zero7	1.00^{+}			
zero1	1.00^{+}			
zero4	1.00^{+}			
zero11	1.00^{+}			
zero5	1.00^{+}			
zero9	1.00^{+}			
lr1	1.00^{+}			
lr2	1.00^{+}			
lr3	1.00^{+}			
lr4	1.00^{+}			
lr5	1.00^{+}			
al1	1.00^{+}			
al2	1.00^{+}			
al3	1.00^{+}			
al4	1.00^{+}			
al5	1.00^{+}			
		Intercepts		
zero7	0.00^{+}			
zero1	0.00^{+}			
zero4	0.00^{+}			
zero11	0.00^{+}			
zero5	0.00^{+}			
zero9	0.00^{+}			
lr1	0.00^{+}			
lr2	0.00^{+}			
lr3	0.00^{+}			
lr4	0.00^{+}			
lr5	0.00^{+}			
al1	0.00^{+}			
al2	0.00^{+}			
al3	0.00^{+}			
al4	0.00^{+}			
al5	0.00^{+}			
		Residual Variances		
zero7	0.46^{+}			

Table D4: Zero-Sum OCFA2

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
Z				
zero7	0.69	0.03	22.91	.000
zero1	0.49	0.03	15.94	.000
zero4	0.61	0.03	20.23	.000
zero11	-0.54	0.03	-18.36	.000
zero5	-0.55	0.03	-18.23	.000
zero9	-0.70	0.03	-23.47	.000
LeftCorrected				
lr1	0.78	0.01	90.67	.000
lr2	0.86	0.01	107.30	.000
lr3	0.90	0.01	119.32	.000
lr4	0.85	0.01	104.22	.000
lr5	0.68	0.01	72.40	.000
AuthCorrected				
al1	0.81	0.01	68.91	.000
al2	0.74	0.01	57.36	.000
al3	0.76	0.01	61.80	.000
al4	0.49	0.01	33.66	.000
al5	0.78	0.01	61.35	.000
Acq				
zero7	0.35	0.01	68.16	.000
zero1	0.35	0.01	68.16	.000
zero4	0.35	0.01	68.16	.000
zero11	0.35	0.01	68.16	.000
zero5	0.35	0.01	68.16	.000
zero9	0.35	0.01	68.16	.000
lr1	0.35	0.01	68.16	.000
lr2	0.35	0.01	68.16	.000
lr3	0.35	0.01	68.16	.000
lr4	0.35	0.01	68.16	.000
lr5	0.35	0.01	68.16	.000
al1	0.35	0.01	68.16	.000
al2	0.35	0.01	68.16	.000
al3	0.35	0.01	68.16	.000
al4	0.35	0.01	68.16	.000
al5	0.35	0.01	68.16	.000
		Intercepts		
zero7	0.00^{+}			
zero1	0.00^{+}			
zero4	0.00^{+}			
zero11	0.00^{+}			
zero5	0.00^{+}			
zero9	0.00^{+}			
lr1	0.00^{+}			
lr2	0.00^{+}			
lr3	0.00^{+}			
lr4	0.00^{+}			
lr5	0.00^{+}			
al1	0.00^{+}			
al2	0.00^{+}			
al3	0.00^{+}			
al4	0.00^{+}			
al5	0.00^{+}			
		Residual Variances		
zero7	0.44^{+}			
zero1	0.60^{+}			

Table D5: Empathy CFA1

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
E				
em1	0.29	0.01	25.79	.000
em2	0.31	0.01	29.67	.000
em3	0.29	0.01	28.58	.000
em4	-0.34	0.01	-26.20	.000
em5	0.28	0.01	28.30	.000
em6	0.25	0.01	21.12	.000
em7	-0.46	0.01	-43.62	.000
em8	-0.49	0.01	-46.32	.000
em9	-0.40	0.01	-29.70	.000
em10	-0.48	0.01	-43.60	.000
LeftCorrected				
lr1	0.83	0.02	51.10	.000
lr2	0.70	0.01	47.49	.000
lr3	0.84	0.01	61.91	.000
lr4	0.82	0.01	56.43	.000
lr5	0.65	0.02	39.65	.000
AuthCorrected				
al1	0.90	0.02	56.38	.000
al2	1.03	0.02	56.25	.000
al3	0.77	0.02	49.03	.000
al4	0.63	0.02	33.54	.000
al5	0.77	0.01	52.97	.000
Acq				
em1	1.00^{+}			
em2	1.00^{+}			
em3	1.00^{+}			
em4	1.00^{+}			
em5	1.00^{+}			
em6	1.00^{+}			
em7	1.00^{+}			
em8	1.00^{+}			
em9	1.00^{+}			
em10	1.00^{+}			
lr1	1.00^{+}			
lr2	1.00^{+}			
lr3	1.00^{+}			
lr4	1.00^{+}			
lr5	1.00^{+}			
al1	1.00^{+}			
al2	1.00^{+}			
al3	1.00^{+}			
al4	1.00^{+}			
al5	1.00^{+}			
		Residual Variances		
em1	0.17	0.01	22.58	.000
em2	0.10	0.01	17.99	.000
em3	0.11	0.01	19.45	.000
em4	0.43	0.01	30.41	.000
em5	0.21	0.01	25.68	.000
em6	0.36	0.01	35.58	.000
em7	0.15	0.01	19.33	.000
em8	0.14	0.01	20.62	.000
em9	0.35	0.01	29.53	.000
em10	0.15	0.01	21.15	.000
lr1	0.74	0.02	34.37	.000

Table D6: Empathy CFA2

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
E				
em1	0.13	0.05	2.85	.004
em2	0.16	0.05	3.43	.001
em3	0.14	0.05	3.00	.003
em4	-0.49	0.05	-10.51	.000
em5	0.13	0.05	2.78	.005
em6	0.10	0.05	2.03	.042
em7	-0.61	0.05	-13.06	.000
em8	-0.64	0.05	-13.47	.000
em9	-0.55	0.05	-11.47	.000
em10	-0.63	0.05	-13.40	.000
LeftCorrected				
lr1	0.85	0.02	50.27	.000
lr2	0.72	0.02	42.60	.000
lr3	0.85	0.01	59.79	.000
lr4	0.83	0.02	52.56	.000
lr5	0.67	0.02	32.38	.000
AuthCorrected				
al1	0.92	0.02	54.68	.000
al2	1.08	0.03	40.85	.000
al3	0.79	0.02	47.11	.000
al4	0.65	0.02	32.27	.000
al5	0.80	0.02	45.73	.000
Acq				
em1	0.27	0.03	10.11	.000
em2	0.27	0.03	10.11	.000
em3	0.27	0.03	10.11	.000
em4	0.27	0.03	10.11	.000
em5	0.27	0.03	10.11	.000
em6	0.27	0.03	10.11	.000
em7	0.27	0.03	10.11	.000
em8	0.27	0.03	10.11	.000
em9	0.27	0.03	10.11	.000
em10	0.27	0.03	10.11	.000
lr1	0.27	0.03	10.11	.000
lr2	0.27	0.03	10.11	.000
lr3	0.27	0.03	10.11	.000
lr4	0.27	0.03	10.11	.000
lr5	0.27	0.03	10.11	.000
al1	0.27	0.03	10.11	.000
al2	0.27	0.03	10.11	.000
al3	0.27	0.03	10.11	.000
al4	0.27	0.03	10.11	.000
al5	0.27	0.03	10.11	.000
		Residual Variances		
em1	0.16	0.01	22.49	.000
em2	0.10	0.01	17.79	.000
em3	0.10	0.01	19.17	.000
em4	0.44	0.01	30.51	.000
em5	0.22	0.01	25.49	.000
em6	0.37	0.01	35.21	.000
em7	0.15	0.01	19.29	.000
em8	0.14	0.01	20.39	.000
em9	0.34	0.01	29.51	.000
em10	0.15	0.01	21.03	.000
lr1	0.74	0.02	[illegible]	.000

Table D7: Empathy OCFA1

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
E				
em1	0.61	0.01	57.76	.000
em2	0.69	0.01	72.58	.000
em3	0.66	0.01	67.06	.000
em4	-0.53	0.01	-51.06	.000
em5	0.63	0.01	58.31	.000
em6	0.48	0.01	40.52	.000
em7	-0.77	0.01	-96.82	.000
em8	-0.78	0.01	-101.73	.000
em9	-0.56	0.01	-52.87	.000
em10	-0.78	0.01	-101.31	.000
LeftCorrected				
lr1	0.68	0.01	79.26	.000
lr2	0.80	0.01	121.58	.000
lr3	0.85	0.01	146.89	.000
lr4	0.81	0.01	122.76	.000
lr5	0.68	0.01	79.69	.000
AuthCorrected				
al1	0.81	0.01	110.46	.000
al2	0.73	0.01	80.73	.000
al3	0.76	0.01	94.12	.000
al4	0.50	0.01	41.70	.000
al5	0.81	0.01	109.93	.000
Acq				
em1	1.00^{+}			
em2	1.00^{+}			
em3	1.00^{+}			
em4	1.00^{+}			
em5	1.00^{+}			
em6	1.00^{+}			
em7	1.00^{+}			
em8	1.00^{+}			
em9	1.00^{+}			
em10	1.00^{+}			
lr1	1.00^{+}			
lr2	1.00^{+}			
lr3	1.00^{+}			
lr4	1.00^{+}			
lr5	1.00^{+}			
al1	1.00^{+}			
al2	1.00^{+}			
al3	1.00^{+}			
al4	1.00^{+}			
al5	1.00^{+}			
		Intercepts		
em1	0.00^{+}			
em2	0.00^{+}			
em3	0.00^{+}			
em4	0.00^{+}			
em5	0.00^{+}			
em6	0.00^{+}			
em7	0.00^{+}			
em8	0.00^{+}			
em9	0.00^{+}			
em10	0.00^{+}			
lr1	0.00^{+}			

Table D8: Empathy OCFA2

	Model			
	Estimate	Std. Err.	z	p
		Factor Loadings		
E				
em1	0.65	0.04	17.94	.000
em2	0.73	0.04	20.26	.000
em3	0.71	0.04	19.51	.000
em4	-0.49	0.04	-13.76	.000
em5	0.68	0.04	18.83	.000
em6	0.53	0.04	14.65	.000
em7	-0.73	0.04	-20.35	.000
em8	-0.74	0.04	-20.68	.000
em9	-0.52	0.04	-14.40	.000
em10	-0.74	0.04	-20.52	.000
LeftCorrected				
lr1	0.79	0.01	63.09	.000
lr2	0.88	0.01	73.18	.000
lr3	0.95	0.01	88.49	.000
lr4	0.89	0.01	76.14	.000
lr5	0.73	0.01	53.14	.000
AuthCorrected				
al1	0.87	0.01	85.17	.000
al2	0.75	0.01	64.20	.000
al3	0.83	0.01	76.91	.000
al4	0.59	0.01	42.09	.000
al5	0.87	0.01	83.33	.000
Acq				
em1	0.33	0.01	49.43	.000
em2	0.33	0.01	49.43	.000
em3	0.33	0.01	49.43	.000
em4	0.33	0.01	49.43	.000
em5	0.33	0.01	49.43	.000
em6	0.33	0.01	49.43	.000
em7	0.33	0.01	49.43	.000
em8	0.33	0.01	49.43	.000
em9	0.33	0.01	49.43	.000
em10	0.33	0.01	49.43	.000
lr1	0.33	0.01	49.43	.000
lr2	0.33	0.01	49.43	.000
lr3	0.33	0.01	49.43	.000
lr4	0.33	0.01	49.43	.000
lr5	0.33	0.01	49.43	.000
al1	0.33	0.01	49.43	.000
al2	0.33	0.01	49.43	.000
al3	0.33	0.01	49.43	.000
al4	0.33	0.01	49.43	.000
al5	0.33	0.01	49.43	.000
		Intercepts		
em1	0.00^{+}			
em2	0.00^{+}			
em3	0.00^{+}			
em4	0.00^{+}			
em5	0.00^{+}			
em6	0.00^{+}			
em7	390.00^{+}			
em8	0.00^{+}			
em9	0.00^{+}			
em10	0.00^{+}			
lr1	0.00^{+}			