Appendix

Online Supplementary Materials for
Patterns of individual differences in the perception of
missing-fundamental tones

D. Robert Ladd[1], Rory Turnbull[1], Charlotte Browne[1], Catherine Caldwell-Harris[2], Lesya Ganushchak[3], Kate Swoboda[1], Verity Woodfield[1], and Dan Dediu[3]

[1]School of Philosophy, Psychology and Language Sciences, University of Edinburgh
[2]Department of Psychology, Boston University
[3]Max-Planck-Institute for Psycholinguistics, Nijmegen

Corresponding author: Prof. D. Robert Ladd, School of Philosophy, Psychology and Language Sciences, University of Edinburgh, 3 Charles Street, Edinburgh EH8 9AD, Scotland. Email bob.ladd@ed.ac.uk

Summary

This appendix forms part of the Supplementary Online Materials and contains more details about the stimuli and the statistical analyses and results described in the main paper.

The Stimuli

Sound files of selected stimuli are provided here, arranged so that readers can easily make some of the comparisons discussed in the main paper. Except where indicated,

the stimuli illustrated here have phase-controlled partials (see 'Stimulus preparation' in the main paper).

The stimuli are identified here by code numbers consisting of three segments, e.g. 56.1600.BA. The first segment refers in abbreviated form to the spectral composition; the second gives the top frequency in Hz; and the third specifies the order of the two component tones (AB or BA). The code for spectral composition indicates the highest harmonics in Tone A and Tone B: thus '56' refers to stimuli in which Tone A consists of harmonics 3, 4 and 5 and Tone B consists of harmonics 4, 5 and 6. For more detail see Table 1 in the main paper.

'Up' responses indicate a spectral percept with AB stimuli and an F0 percept with BA stimuli. 'Down' responses indicate a spectral percept with BA stimuli and an F0 percept with AB stimuli.

Order. Here are two pairs of stimuli that are identical except for order:

45.0675.AB 45.0675.BA
67.1600.AB 67.1600.BA

Spectral composition. Here are two triplets of stimuli that are identical except for spectral composition:

45.1200.AB 56.1200.AB 67.1200.AB
45.2150.BA 56.2150.BA 67.2150.BA

These are 'identical' only in the context of our stimulus matrix, i.e. in the sense that the top frequency and the order are the same. The missing fundamentals, and the intervals between them, are different in each case, because of the interdependencies discussed in the section 'Stimulus variables in the MF task' in the main paper.

Top frequency. Here is a set of stimuli that are identical except for top frequency:
45.0500.AB 45.0675.AB 45.0900.AB 45.1200.AB 45.1600.AB 45.2150.AB
Some readers may observe that their response changes from 'up' to 'down' as the top frequency increases. This is the pattern of responses illustrated in Figure 3 of the main paper and may be related to the order effect illustrated in Figure 4 of the main paper.

Phase. Here is a pair of stimuli that differ only in whether phase was controlled in generating them:

56.0500.BA (phase controlled)

nophase.56.0500.BA (phase not controlled)

Duration. Here are two pairs (from the test-retest material in Experiment 4c) that differ only in duration:

67.0900.AB (500ms tones, 250ms gap) short.67.0900.AB (180ms tones, 20ms gap)

45.0500.BA (500ms tones, 250ms gap) short.45.0500.BA (180ms tones, 20ms gap)

The Schneider Index

As described in the main paper, each stimulus is described by a *Top Frequency Level* (here denoted *FL*), a *Spectral Composition* (denoted *SC*), and a *Direction* (AB or BA) and any single participant gives an "Up" or "Down" response for each such stimulus. By definition, we created the stimuli in such a way that a "Down" response to the AB order and an "Up" response to the BA order are based on the missing fundamental (*F0* responses), while "Up" for AB and "Down" for BA are based on the harmonics present in the stimuli (*spectral* responses). See Figure 1 and associated discussion in the main paper for more details

For a set of stimuli and responses, we can count the number of F0 (denoted f0) and spectral (denoted sp) responses; with these, the *Schneider Index* (denoted SI) for this set of stimuli is:

$$SI = \frac{sp - f0}{sp + f0} \quad (1)$$

which can vary between –1 (100% F0 responses) and +1 (100% spectral responses). When considering *all* the responses given by the participant in all FL and SC conditions and both orders, we compute the *overall Schneider Index*, denoted here as SI_O, but we can also compute *partial Schneider Indices* by restricting the value of some experimental parameters, such as for a given FL value (say, 500; denoted $SI_{FL=500}$), a given SC value (say, using the notations defined above, 45; $SI_{SC=45}$), or for a single cell in the FL×SC stimulus matrix (say FL=500 and SC=45; $SI_{FL=500,SC=45}$), or even for a given order (such as SI_{AB}). While measures comparable to the overall SI_O have been used in the previous literature as a measure of a participant's global style,

the various partial SI allow us a better understanding of the individual differences in the perception of the missing fundamental and the factors that affect them.

Summaries based on the partial Schneider Indices

Using various partial SI measures, we can define a number of summaries capturing different aspects of the participants' behaviour. One such summary captures the magnitude of the *order effect*, quantifying the difference that the two orders of presentation (AB and BA) might have on the participants' answers. We will denote this as *OE*, defined as the mean absolute difference between the partial SI for the AB and BA orders:

$$\mathrm{OE} = \underset{fl,sc}{mean} \frac{|\,\mathrm{SI}_{\mathrm{FL}=fl,\mathrm{SC}=sc,\mathrm{AB}} - \mathrm{SI}_{\mathrm{FL}=fl,\mathrm{SC}=sc,\mathrm{BA}}\,|}{2} \tag{2}$$

where *fl* is a valid FL value, *sc* is a valid SC value, $|x|$ is the absolute value (modulus) of x, and division by 2 insures that OE lies between 0 (no order effect), and 1 (maximum possible order effect, i.e. consistently opposite responses to AB and BA items).

Other summaries are driven by the structure of the data, in the sense that they are derived from the principal components in a *Principal Components Analysis* (*PCA*; Jolliffe 2002). PCA is a technique which transforms a set of N inter-correlated variables into the same number of independent components ordered by the amount of variation in the data they explain. Thus, the first component, PC_1, explains most of the variation in the data, followed by PC_2, and so on. Given that these components are linear combinations of the original variables weighted by their loadings, these loadings can be used to interpret the meaning of the components. We performed PCA on the cell-level SI as the N inter-correlated variables: $\mathrm{SI}_{\mathrm{FL}=fl,\mathrm{SC}=sc}$, where *fl* and *sc* represent valid values of FL and SC as above.

As described in the main paper, we found that the first and second components, PC_1 and PC_2, are extremely similar across experiments in both structure (i.e., the loadings of the cell-level SI) and the amount of variance explained (Table S1). More precisely, PC_1 explains about half of the variance and is equivalent to the overall SI (SI_O) in the sense that all cell-level SI have loadings of same sign and

comparable magnitude. Therefore, we abstracted away from the actual loadings and defined PC_1^* as

$$PC_1^* = \underset{fl,sc}{mean}\left(SI_{FL=fl,SC=sc}\right) \tag{3}$$

whereby we assign the same weight to all FL×SC cells. As expected, PC_1^* is highly correlated with SI_O (overall SI) ($r = .99$, $p < .001$). In the main paper we therefore used SI_O (there denoted simply SI) rather than PC_1^*.

PC_2 expresses one of the response patterns we identified for participants with intermediate SI_O values, namely a difference between responses at lower and higher FL. Specifically, the loadings of the cell-level SI of the cells with FL lower than a threshold, *T*, have the opposite sign to those of the cells with FL higher than *T*. This is the pattern of responses illustrated in Figure 3 of the main paper. We interpreted this pattern as reflecting a type of consistency in the participant's responses and, as above, we abstracted away from the actual loadings by formalizing it as the *Consistency Index* (denoted *CI*) as follows:

$$CI = \frac{\underset{fl<T,sc}{mean}\left(SI_{FL=fl,SC=sc}\right) - \underset{fl\geq T,sc}{mean}\left(SI_{FL=fl,SC=sc}\right)}{2} \tag{4}$$

In practice, the threshold *T* depends on the actual experiment (its approximate values are 1000Hz, 1200Hz, 1600Hz, 1800Hz, 1000Hz, 1000Hz and 1000Hz for experiments 1, 2, 3a, 3b, 4a, 4b, and 4c, respectively), but we were able to find a general formula relating it to the limits of the FL space. We did this by regressing the observed *T* values on the minimum and maximum FL values across the experiments, as we found that:

$$T^* = 800 + \frac{FL_{max} - FL_{min}}{5} \tag{5}$$

gives a good approximation for the observed threshold values, *T*. CI expresses the size of the difference between low and high FL, with strongly polarised values (near –1.0 and +1.0) representing a large difference, and values near 0.0 representing a small difference. Strongly positive values indicate a switch from spectral to F0 responses as

FL increases; strongly negative values would mean the reverse, but as noted in the main paper strongly negative values are never found. Among participants with strongly positive CI, the location of the crossover point (the threshold T^*) on the FL dimension is fairly consistent, somewhere in the vicinity of 1000 Hz.

Table S1
Principal Components Analysis (PCA) for each experiment.

Experiment Cell (SC, FL)	**PC_1**	**PC_2**
Experiment 1		
Eigenvalues: **PC_1 (3.79)**, **PC_2 (1.16)**, PC_3 (0.45)		
Variance explained	***56.99%***	***17.39%***
56, 300	*-0.27*	*-0.25*
56, 500	*-0.28*	*-0.29*
56, 900	*-0.30*	0.02
56, 1400	*-0.22*	**0.34**
56, 2200	*-0.17*	**0.19**
78, 300	*-0.24*	*-0.32*
78, 500	*-0.31*	*-0.32*
78, 900	*-0.34*	0.02
78, 1400	*-0.26*	**0.36**
78, 2200	*-0.17*	**0.24**
90, 300	*-0.13*	*-0.22*
90, 500	*-0.33*	*-0.23*
90, 900	*-0.31*	0.09
90, 1400	*-0.24*	**0.37**
90, 2200	*-0.19*	**0.28**
Experiment 2		
Eigenvalues: **PC_1 (3.55)**, PC_2 (0.47), PC_3 (0.14)		
Variance explained	***80.31%***	***10.67%***
56, 500	*-0.30*	**0.47**
56, 750	*-0.37*	**0.21**
56, 1050	*-0.34*	**0.10**
56, 1400	*-0.35*	-0.04
56, 1800	*-0.29*	*-0.36*
89, 500	*-0.32*	**0.49**
89, 750	0.00	0.05
89, 1050	*-0.35*	-0.05
89, 1400	*-0.34*	*-0.41*
89, 1800	*-0.32*	*-0.43*
Experiment 3a		
Eigenvalues: **PC_1 (10.02)**, **PC_2 (3.24)**, **PC_3 (1.06)**		
Variance explained	***56.84%***	***18.36%***
56, 250	-0.03	-0.08
56, 500	*-0.12*	*-0.22*
56, 750	**0.18**	*-0.31*
56, 1050	*-0.21*	*-0.25*
56, 1400	*-0.20*	*-0.22*

56, 1800	*-0.20*	0.02
56, 2000	*-0.24*	0.01
56, 2500	*-0.21*	**0.19**
56, 3000	*-0.19*	**0.12**
56, 4000	*-0.24*	**0.22**
56, 5000	*-0.23*	**0.26**
56, 6000	*-0.13*	**0.31**
78, 300	0.02	*-0.15*
78, 500	-0.07	*-0.18*
78, 900	*-0.22*	*-0.28*
78, 1400	*-0.22*	*-0.15*
78, 2200	*-0.20*	**0.12**
89, 250	0.01	*-0.11*
89, 500	-0.06	*-0.15*
89, 750	*-0.14*	*-0.30*
89, 1050	*-0.23*	*-0.22*
89, 1400	*-0.23*	-0.06
89, 1800	*-0.25*	0.03
89, 2000	*-0.24*	0.07
89, 2500	*-0.23*	**0.11**
89, 3000	*-0.24*	**0.11**
89, 4000	*-0.19*	**0.18**
89, 5000	*-0.17*	**0.23**
89, 6000	-0.06	**0.10**
Experiment 3b		
Eigenvalues: **PC_1 (3.90)**, **PC_2 (1.13)**, PC_3 (0.55)		
Variance explained	***46.50%***	***13.46%***
56, 250	**0.12**	-0.09
56, 300	**0.18**	-0.04
56, 400	**0.20**	**0.13**
56, 500	**0.24**	0.07
56, 700	**0.26**	**0.24**
56, 900	**0.26**	**0.13**
56, 1150	**0.24**	0.09
56, 1400	**0.24**	0.06
56, 1700	**0.22**	0.00
56, 2000	**0.19**	*-0.20*
56, 2500	**0.16**	*-0.25*
56, 3000	**0.14**	*-0.27*
56, 4000	**0.13**	*-0.31*
56, 5000	**0.13**	*-0.34*
89, 250	-0.01	0.06
89, 300	0.05	-0.02
89, 400	**0.15**	**0.14**
89, 500	**0.18**	**0.20**
89, 700	**0.24**	**0.23**
89, 900	**0.23**	**0.19**
89, 1150	**0.25**	**0.19**
89, 1400	**0.24**	0.05
89, 1700	**0.24**	-0.03
89, 2000	**0.19**	*-0.19*
89, 2500	**0.14**	*-0.20*
89, 3000	**0.14**	*-0.31*
89, 4000	**0.11**	*-0.29*
89, 5000	**0.11**	*-0.23*
Experiment 4a		
Eigenvalues: **PC_1 (3.37)**, PC_2 (0.67), PC_3 (0.26)		

Variance explained	***61.29%***	***12.10%***
45, 500	**0.16**	**0.42**
45, 675	**0.20**	**0.37**
45, 900	**0.31**	0.07
45, 1200	**0.27**	*-0.11*
45, 1600	**0.30**	*-0.23*
45, 2150	**0.29**	*-0.22*
56, 500	**0.17**	**0.33**
56, 675	**0.21**	**0.20**
56, 900	**0.28**	0.08
56, 1200	**0.23**	*-0.20*
56, 1600	**0.23**	*-0.16*
56, 2150	**0.21**	*-0.21*
67, 500	**0.12**	**0.43**
67, 675	**0.16**	**0.26**
67, 900	**0.24**	-0.01
67, 1200	**0.22**	-0.09
67, 1600	**0.27**	*-0.17*
67, 2150	**0.27**	*-0.08*
Experiment 4a (retest)		
Eigenvalues: **PC_1 (3.64)**, PC_2 (0.98), PC_3 (0.33)		
Variance explained	***58.60%***	***15.83%***
45, 500	**0.17**	**0.40**
45, 675	**0.24**	**0.35**
45, 900	**0.29**	-0.06
45, 1200	**0.31**	*-0.14*
45, 1600	**0.25**	*-0.18*
45, 2150	**0.27**	*-0.23*
56, 500	**0.19**	**0.25**
56, 675	**0.19**	**0.26**
56, 900	**0.28**	0.09
56, 1200	**0.24**	*-0.14*
56, 1600	**0.25**	*-0.12*
56, 2150	**0.19**	*-0.23*
67, 500	**0.10**	**0.49**
67, 675	**0.20**	**0.27**
67, 900	**0.26**	0.02
67, 1200	**0.23**	-0.09
67, 1600	**0.24**	*-0.10*
67, 2150	**0.25**	*-0.23*
Experiment 4b		
Eigenvalues: **PC_1 (6.40)**, PC_2 (0.52), PC_3 (0.21)		
Variance explained	***79.59%***	***6.50%***
45, 500	*-0.17*	**0.43**
45, 675	*-0.25*	**0.26**
45, 900	*-0.25*	**0.15**
45, 1200	*-0.28*	-0.03
45, 1600	*-0.27*	*-0.18*
45, 2150	*-0.24*	*-0.28*
56, 500	*-0.20*	**0.22**
56, 675	*-0.24*	**0.14**
56, 900	*-0.26*	0.08
56, 1200	*-0.25*	*-0.16*
56, 1600	*-0.25*	*-0.20*
56, 2150	*-0.25*	*-0.24*
67, 500	*-0.09*	**0.33**
67, 675	*-0.18*	**0.22**

67, 900	*-0.24*	**0.26**
67, 1200	*-0.22*	0.09
67, 1600	*-0.25*	*-0.14*
67, 2150	*-0.29*	*-0.41*
Experiment 4c		
Eigenvalues: **PC_1 (7.01)**, PC_2 (0.70), PC_3 (0.32)		
Variance explained	***75.91%***	***7.53%***
45, 500	*-0.23*	*-0.36*
45, 675	*-0.27*	*-0.22*
45, 900	*-0.27*	0.05
45, 1200	*-0.27*	**0.13**
45, 1600	*-0.26*	**0.24**
45, 2150	*-0.22*	**0.25**
56, 500	*-0.20*	*-0.38*
56, 675	*-0.20*	*-0.36*
56, 900	*-0.24*	*-0.10*
56, 1200	*-0.25*	**0.14**
56, 1600	*-0.23*	**0.24**
56, 2150	*-0.22*	**0.23**
67, 500	*-0.15*	*-0.30*
67, 675	*-0.20*	*-0.17*
67, 900	*-0.24*	*-0.18*
67, 1200	*-0.26*	0.02
67, 1600	*-0.26*	**0.21**
67, 2150	*-0.25*	**0.28**
Experiment 4c (retest)		
Eigenvalues: **PC_1 (6.13)**, PC_2 (0.55), PC_3 (0.37)		
Variance explained	***75.96%***	***6.79%***
45, 500	*-0.24*	*-0.27*
45, 675	*-0.26*	*-0.16*
45, 900	*-0.27*	**0.15**
45, 1200	*-0.26*	**0.11**
45, 1600	*-0.25*	**0.27**
45, 2150	*-0.21*	**0.29**
56, 500	*-0.19*	*-0.40*
56, 675	*-0.19*	*-0.29*
56, 900	*-0.26*	*-0.23*
56, 1200	*-0.25*	**0.14**
56, 1600	*-0.26*	**0.21**
56, 2150	*-0.20*	**0.25**
67, 500	*-0.18*	*-0.36*
67, 675	*-0.22*	*-0.25*
67, 900	*-0.25*	*-0.16*
67, 1200	*-0.24*	**0.11**
67, 1600	*-0.26*	**0.15**
67, 2150	*-0.24*	**0.20**

Note. Results from Principal Component Analysis (PCA) for each experiment separately showing the first two PCs, the variance they explain, the eigenvalues for the first three PCs (those greater than 1 are in bold), and the loadings of the cell-level SI are shown. The signs on the loadings are arbitrary but their pattern of contrast is not; we use italic to highlight negative loadings and bold to highlight positive loadings, with an arbitrary threshold of ±0.10 for difference from 0 (regular font). Note that in these analyses we also include the retest data from Experiments 4a and 4c, treating them as separate experiments referred to here as "4a (retest)" and "4c (retest)".

Finding clusters of similar participants

In summary, we were able to provide a good characterization of participants' behaviour on the missing fundamental task using three summaries: the overall Schneider Index (SI), the Consistency Index (CI), and the order effect (OE). Each participant's behaviour can therefore be conceptualized as a point in a 3-dimensional space defined by SI×CI×OE and bounded between –1 and +1 on the first dimension, –1 and +1 on the second, and 0 and 1 on the third. A visual representation of all our participants is given in Figures 6A and 6B in the main paper, showing the two projections on the SI×CI and SI×OE. These same representations are shown here in color as Figures S1A and S1B.

To ascertain the apparent existence of groups of participants with similar behaviour, we conducted a *k-means clustering* analysis (Hartigan & Wong 1979). For a given number of groups, *k*, this method tries to allocate each participant to the group with the closest mean, based on the inter-participant distances. We computed these

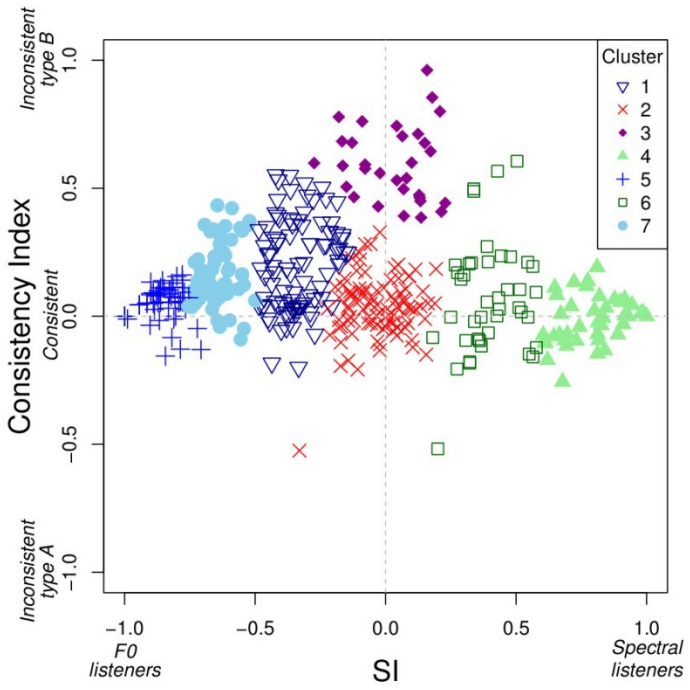

Figure S1A (Color version of Figure 6A in the main paper): Distribution of participants across all experiments in the SI×CI space, also showing the seven

optimal clusters using symbols and colours. 'Inconsistent type A' participants give predominantly spectral responses to lower frequency stimuli and F0 responses to higher frequency stimuli; the 'Inconsistent type B' pattern of responses would be the reverse, but as can be seen this pattern is not found.

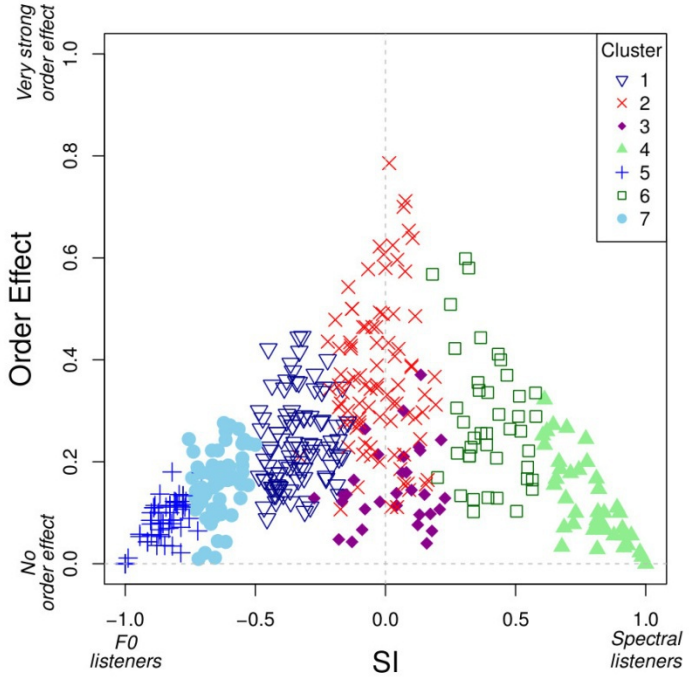

Figure S1B (Color version of Figure 6B in the main paper): Distribution of participants across all experiments in the SI×OE space, also showing the seven optimal clusters using symbols and colors.

distances as the Euclidean distances between all pairs of participants in the 3-dimensional space SI×CI×OE; therefore participants with very similar behaviour will have very small distances between them, while participants who differ will be further away. However, we do not know *a priori* the optimum number of such clusters, k, and therefore we used an automatic search procedure which selects the best k on the basis of the *Calinski Harabasz index* (Calinski & Harabasz 1974; function `kmeansruns` in `R`'s library `fpc`), which optimizes the within- versus between-cluster distances.

We found that the optimal value of k is 7, and the clusters, as discussed in the main paper, are quite interpretable. To test the robustness of k=7, we noted that the Euclidean distance is a particular case (with p=2) of the general Minkowski distances, which for a pair of n-dimensional points $(x_i)_{i=1,n}$ and $(y_i)_{i=1,n}$ is defined as $(\sum_{i=1,n} |x_i - y_i|^p)^{1/p}$, where the order p is fixed. Thus, we repeatedly computed the optimal number of clusters k for different orders p, and we found that k=7 is robust for $2 \le p \le 6$, while

for Manhattan distances (p=1), k=3, and for Euclidean distances (p=2) k=10 is equally good and suggest very similar clusters. The k=7 clusters obtained using the Euclidean distances are shown in Figure S1 (panels A and B).

Combining different experiments

As described in the main paper, we amalgamated the results from several experiments conducted at different times and in different places, with different participant groups and slightly different stimulus characteristics. This necessarily raises the question of the legitimacy of this procedure. Here we expand on the justifications given in the main paper.

As described in detail above, we conducted separate Principal Component Analyses for each experiment before amalgamating them. For all experiments we found similar patterns reflected by the first two principal components, not only in terms of structure but also in the amount of variance explained. This provides *a priori* support for the idea that the experiments reveal fundamentally similar patterns of behaviour. Impressionistic comparison of the distributions of the three data summaries derived above (SI, CI and OE) for the separate experiments also seems to confirm that the results of the individual experiments are quite similar, with the possible exception of 3a and 3b. These comparisons are illustrated in Figure S2 (panels A, B and C), in which the distributions are smoothed using Kernel Density Estimation (cf. Figure 5 in the main paper).

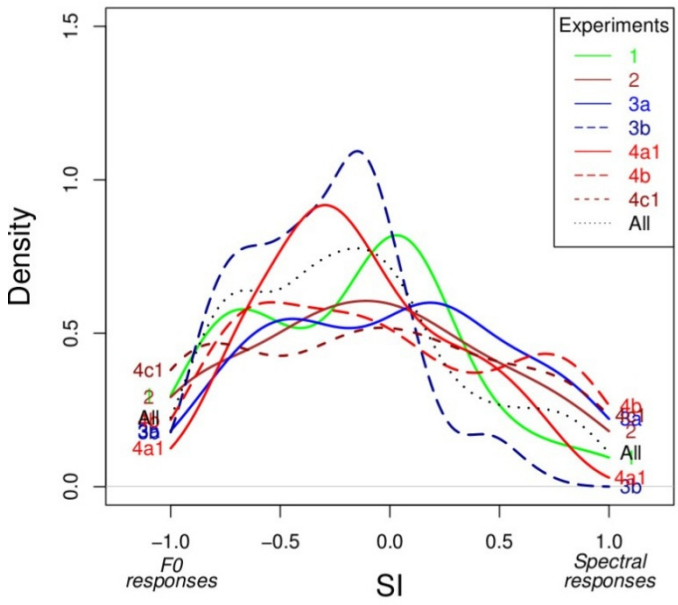

Figure S2A: Smoothed distribution of SI (overall Schneider Index) across experiments.

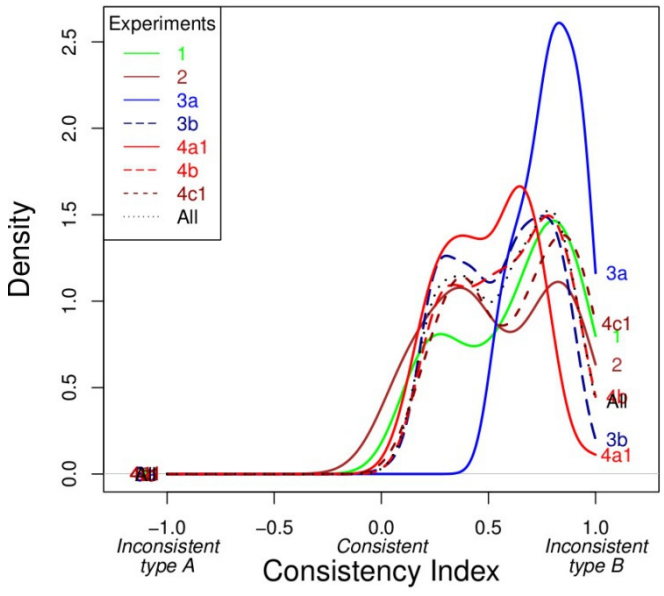

Figure S2B: Smoothed distribution of CI (Consistency Index) across experiments.

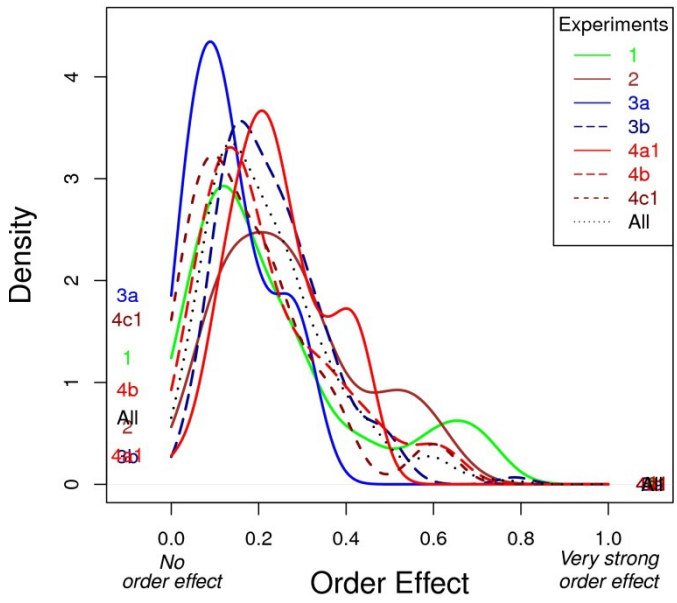

Figure S2C**:** Smoothed distribution of OE (Order Effect) across experiments.

To test more rigorously for differences between experiments, we conducted ANOVAs followed by pair-wise t-tests (corrected for multiple testing using Tukey's Honest Significant Difference (as implemented in **R** by the `aov` and `TukeyHSD` functions) as well as Kolmogorov-Smirnov tests (implemented in **R** by the `ks.test` function) corrected for multiple testing using Holm's (1979) method. The results of these analyses are summarized in Table S2. Overall, there are few significant differences, confirming that the patterns across the experiments are very similar, and supporting the idea that the task is robust to small differences of methodology. At the same time, these tests confirm that 3a and to a lesser extent 3b are somewhat different, especially with respect to the Consistency Index and the Order Effect. We therefore conducted supplementary clustering analyses, first excluding Experiment 3a and then excluding both Experiments 3a and 3b. These analyses yielded a similar cluster structure, though the actual number of clusters varied: excluding 3a yielded 10 as the optimal number of clusters, while excluding 3b gave an optimum of 4. In both cases, however, 7 was also close to this optimum. These results suggest that the inclusion of Experiments 3a and 3b does not distort the conclusions reported in the main paper based on the amalgamated results.

Table S2

Comparison of individual experiments

Measure	ANOVA		Pair-wise *t*-tests			Kolmogorov-Smirnov	
	***F*(6,405)**	***p***	**Groups**	**Difference**	***p***	**Groups**	***p***
SI	3.96	0.0007	3b-4b	0.27	0.0014	3b-4b	0.0095
						3a-3b	0.017
						3b-4c1	0.0068
CI	4.96	6.53e-05	2-3a	-0.25	0.0088		
			3a-3b	0.23	0.00025	3a-3b	0.008
			3a-4a1	0.29	0.000028	3a-4a1	0.00059
			3a-4b	0.20	0.0075	3a-4b	0.048
OE	3.09	0.0057	2-3a	0.14	0.02		
			3a-3b	-0.09	0.045	3a-3b	0.031
			3a-4a1	-0.11	0.038	3a-4a1	0.039
						3b-4c1	0.041

Note. The pair-wise t-tests and Kolmogorov-Smirnoff tests are corrected for multiple testing using Tukey's HSD and Holm's procedures respectively.

References

Calinski, R. B., & Harabasz, J. (1974). A dendrite method for cluster analysis, *Communications in Statistics*, 3, 1-27. doi:10.1080/03610927408827101

Hartigan, J. A., & Wong, M. A. (1979). A k-means clustering algorithm. *Applied Statistics*, 28, 100–108. doi:10.2307/2346830

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scandinavian Journal of Statistics* 6:65–70.

Jolliffe, I. (2002). *Principal Component Analysis.* Springer Verlag: New York.