

Political Cultures:
Measuring Values Heterogeneity
Supplemental Information

Lisa Blaydes* Justin Grimmer†

March 15, 2019

1 Model

$$\begin{aligned}
 \alpha_k &\sim \text{Gamma}(0.25, 1) \\
 \boldsymbol{\theta}_{r,k} &\sim \text{Dirichlet}(\boldsymbol{\lambda}) \\
 \boldsymbol{\pi}_i &\sim \text{Dirichlet}(\boldsymbol{\alpha}) \\
 \boldsymbol{\tau}_{ij} | \boldsymbol{\pi}_i &\sim \text{Multinomial}(1, \boldsymbol{\pi}_i) \\
 \mathbf{y}_{ijr} | \tau_{ijk} = 1, \boldsymbol{\theta}_{kr} &\sim \text{Multinomial}(1, \boldsymbol{\theta}_{kr})
 \end{aligned}$$

If we assume $\boldsymbol{\lambda} = 1$, then the full posterior is given by,

$$p(\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\tau} | \mathbf{Y}) \propto \prod_{k=1}^K \frac{\exp(-\frac{\alpha_k}{1/4})}{1/4} \times \prod_{i=1}^N \left[\frac{\Gamma(\sum_{k=1}^K \alpha_k)}{\prod_{k=1}^K \Gamma(\alpha_k)} \prod_{k=1}^K \pi_{ik}^{\alpha_k - 1} \prod_{j=1}^{N_i} \prod_{k=1}^K \left[\pi_{ik} \prod_{r=1}^R \left(\prod_{q=1}^{Q_r} \theta_{r k q}^{I(y_{ijr}=q)} \right) \right]^{\tau_{ijk}} \right]$$

2 Estimation

Variational approximations are a deterministic method to approximate a posterior. First, we assume a specific — though still very general — functional form for an approximating distribution. Then, we use an iterative algorithm to find the member of this approximating distribution family that is closest to the true posterior.

We will assume an approximation distribution given by,

$$q(\boldsymbol{\alpha}, \boldsymbol{\theta}, \boldsymbol{\pi}, \boldsymbol{\tau}) = q(\boldsymbol{\alpha}) \prod_{i=1}^N \left[q(\boldsymbol{\pi})_i \prod_{j=1}^{N_i} q(\boldsymbol{\tau})_j \right] \prod_{r=1}^R \prod_{k=1}^K q(\boldsymbol{\theta})_{kr}$$

*Department of Political Science, Stanford University

†Department of Political Science, Stanford University

We use the following algorithm to find the member of this family that is closest to the true posterior (Bishop, 2006; Grimmer, 2011)

2.1 Update for $q(\boldsymbol{\tau})_{ij}$

$q(\boldsymbol{\tau})_{ij}$ is a Multinomial(1, \boldsymbol{r}_{ij}) distribution with typical parameter r_{ijk}

$$r_{ijk} \propto \exp \left\{ \mathbb{E}[\log \pi_{ik}] + \sum_{r=1}^R \left(\sum_{q=1}^{Q_r} y_{ijrp} \mathbb{E}[\log \theta_{rkq}] \right) \right\}$$

We will calculate $\mathbb{E}[\log \pi_{ik}]$ and $\mathbb{E}[\log \theta_{rkq}]$ after deriving the parametric family of their approximating distributions.

2.2 Update for $q(\boldsymbol{\pi})_i$

$q(\boldsymbol{\pi})_i$ is a Dirichlet($\boldsymbol{\gamma}_i$) with typical parameter γ_{ik} ,

$$\gamma_{ik} = \alpha_k + \sum_{j=1}^{N_i} r_{ijk}$$

2.3 Update for $q(\boldsymbol{\theta})_{kr}$

$q(\boldsymbol{\theta})_{kr}$ is a Dirichlet($\boldsymbol{\eta}_{kr}$) distribution, with typical parameter η_{krq} ,

$$\eta_{krq} = \lambda_q + \sum_{i=1}^N \sum_{j=1}^{N_i} r_{ijk} y_{ijrq}$$

2.4 Completing $q(\boldsymbol{\sigma}_i)$ and $q(\boldsymbol{\tau})_{ij}$

To finish the updates, we calculate the following expected values, based on the derived derived distributional forms,

$$\begin{aligned} \mathbb{E}[\log \pi_{ik}] &= \psi(\gamma_{ik}) - \psi\left(\sum_{z=1}^K \gamma_{iz}\right) \\ \mathbb{E}[\log \theta_{krp}] &= \psi(\eta_{krp}) - \psi\left(\sum_{z=1}^K \eta_{krz}\right) \end{aligned}$$

2.4.1 Update Steps for $\boldsymbol{\alpha}$

A closed form update step is unavailable for $\boldsymbol{\alpha}$ so we rely upon an efficient Newton-Raphson algorithm, developed in Minka (2000). Calculating the expected values and differentiating with respect to α_k shows that

$$\frac{\partial \log q(\boldsymbol{\alpha})_k^{\text{new}}}{\partial \alpha_k} = -\frac{1}{\lambda} + N \Psi\left(\sum_{k=1}^K \alpha_k\right) - N \Psi(\alpha_k) + \sum_{i=1}^N \frac{\left(\Psi(\gamma_{itk}) - \Psi\left(\sum_{z=1}^K \gamma_{itz}\right)\right)}{N}$$

and collect all the first derivatives into the (gradient) vector $\frac{\partial \log q(\boldsymbol{\alpha})_k^{\text{new}}}{\partial \boldsymbol{\alpha}_k}$. Define \mathbf{H} as the Hessian (matrix of second derivatives). Typical on diagonal element $h_{jj} = N_s \Psi'(\sum_{k=1}^K \alpha_k) - N \Psi'(\alpha_j)$ where $\Psi'(\cdot)$ is the trigamma function (the derivative of the digamma function) and the typical off-diagonal element ($a \neq b$) we have $h_{ab} = N_z \Psi'(\sum_{k=1}^K \alpha_k)$, so

$$\boldsymbol{\alpha}^{\text{new}} = \boldsymbol{\alpha}^{\text{old}} - \mathbf{H}^{-1} \frac{\partial \log q(\boldsymbol{\alpha})_k^{\text{new}}}{\partial \boldsymbol{\alpha}_k}$$

until the change in $\boldsymbol{\alpha}$ drops below a tolerance level (10^{-8} in the implementation). \mathbf{H} 's structure makes it is easily invertible, making the Newton-Raphson algorithm exceedingly fast.

3 Cross Validation

To select the number of categories we employ 10-fold cross validation (Hastie, Tibshirani and Friedman, 2001) and then use the average predicted probability of question responses. To evaluate model fit using cross validation for each fold of the data we fit our model using a specified number of clusters. For each fit we then compute the maximum a posteriori (MAP) estimates for the country specific distributions of categories. Call this $\hat{\boldsymbol{\pi}}_j$ and call the proportion of respondents in country j w_j with the J length vector of weights \boldsymbol{w} . We then compute the weighted average distribution of sub-cultures across countries as $\tilde{\boldsymbol{\pi}} = \hat{\boldsymbol{\pi}}_j' \boldsymbol{w}$.

For each respondent and each question we then compute a response that averages over the subcultures. We compute the MAP estimate of the subculture specific distributions as $\hat{\boldsymbol{\theta}}_{kr}$. Then, for each question r we compute $\tilde{\boldsymbol{\theta}}_r = \hat{\boldsymbol{\theta}}_{kr}' \tilde{\boldsymbol{\pi}}$. We then assess the predicted probability of the responses for each question and each respondent. Thus our cross validation statistic for k subcultures

$$\text{CV}_k = \sum_{i=1}^N \sum_{j=1}^J \sum_{r=1}^R Y_{ijr} \log \tilde{\boldsymbol{\theta}}_r$$

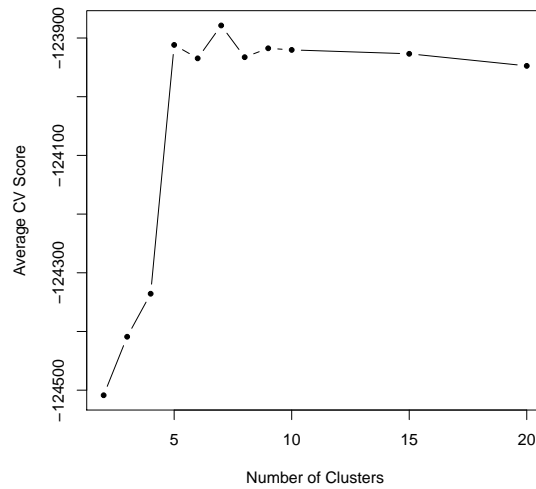
Where we use the natural logarithm to avoid computational underflow. Higher values of the cross validation statistic imply a better fit, because the model is allocating a higher probability to the actual responses.¹ Figure 1 provides the cross validation scores across different numbers of subcultures. This reveals that 7 is the highest score.

4 Validating the Results

Our cross validation procedure provides us with some assurances that our model is fitting the data as well as possible, given the specification. This provides a quantitative evaluation of model fit. But, more importantly, we need to establish that our model is providing substantive insights into the respondent behavior in our data. To establish this, we provide evidence that the types that we identify are describing meaningful variation in the data and we also describe the characteristics of individuals who adopt particular types of subcultures. We also do a ‘‘deep dive’’ on the variation in Figure 1 of the main text and describe country specific variation for a set of countries. Across

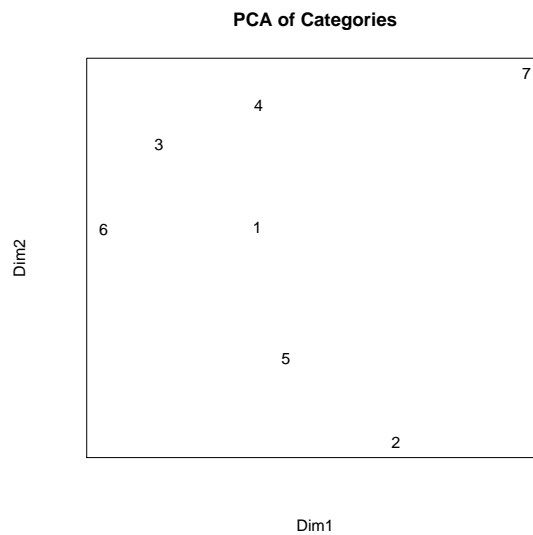
¹Alternatively, we can compute a country specific prediction for each individual, which we discuss in our replication code.

Figure 1: Cross Validation Statistic Across Different Numbers of Clusters



these descriptions, we demonstrate that the method for identifying subcultures yields latent types that provide insights into the underlying culture.

Figure 2: Latent Category PCA



Before discussing specific differences across the categories, Figure 2 provides a PCA of the categories. To do this, we take the first two principal components obtained from trying to explain

the variation of categories across countries. This provides an indication of categories that tend to appear together. As our description suggests, category 6 and 7 are quite distinct on the first dimension. Category 2 and 5 are focused more on material possessions. Overall, this provides further evidence for the description of the categories in the main text.

5 Latent Type Face Validity

In our main text we describe the questions that define what makes the sub-cultures different. In this section we first provide further evidence for the group differences. We do this in two ways.

First, we provide a measure of how distinctive a culture’s response is for each question. To calculate this we assign each respondent to one culture by identifying the most likely category from our model call this $\hat{\tau}_i$. Then, we measure the proportion of respondents from the culture who provided each response, including an additional ”no opinion category”. For culture k and for question r we measure the proportion giving a particular response as $t_{kr} = \frac{\sum_{j=1}^J \sum_{i=1}^{N_j} I(\hat{\tau}_{ijk}=1)I(Y_{ijk}=1)}{\sum_{j=1}^J \sum_{i=1}^{N_j} I(\hat{\tau}_{ijk}=1)}$. (Note that this will be very similar to the MAP estimates of θ_{kr} from the model).

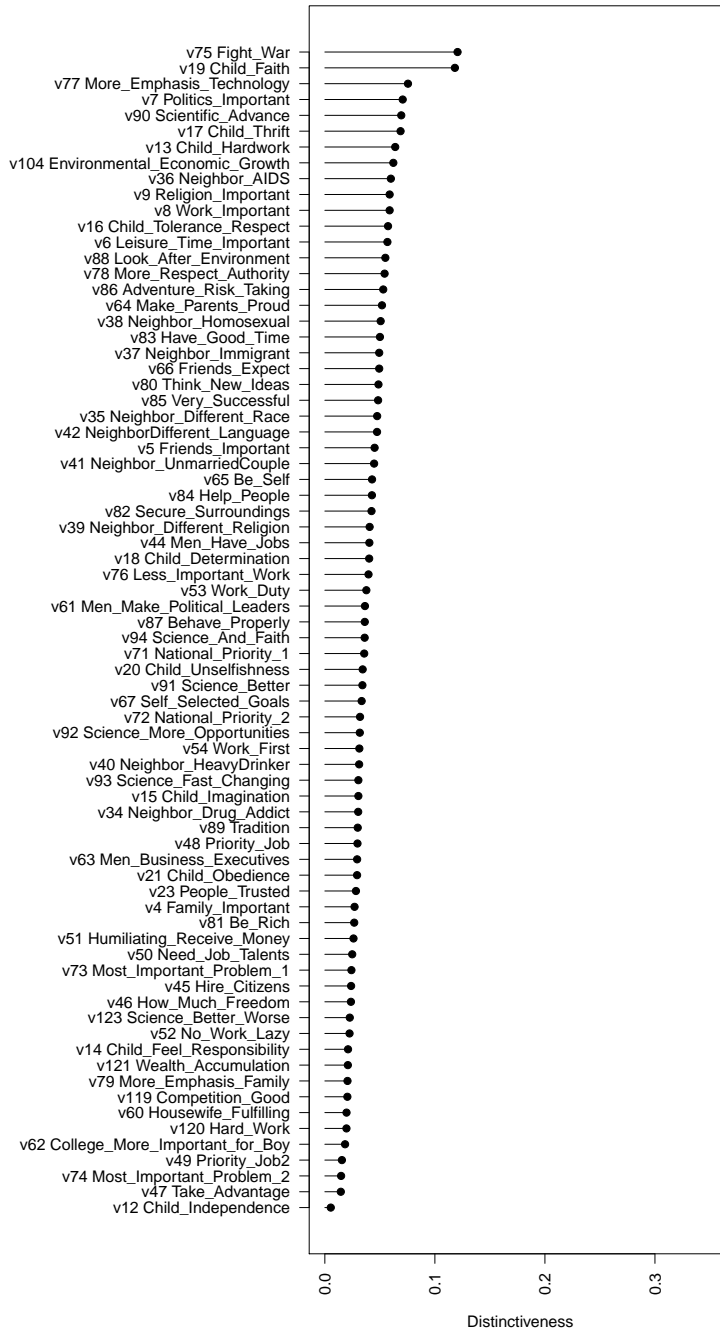
Using this distribution of responses for each culture, we measure the each questions distinctiveness for each culture. To obtain this measure, d_{kr} we compare the distribution of a culture’s responses to the average distribution across the other cultures. Specifically, suppose that for question r there are Q_r potential responses and define the average response for the other $K - 1$ cultures as $t_{-kr} = \sum_{m \neq k} \frac{t_{mr}}{K-1}$. Then we define the distinctiveness for the question as the average difference for each question option:

$$d_{kr} = \sum_{q=1}^{Q_r-1} \frac{|d_{krq} - t_{-krq}|}{Q_r - 1}$$

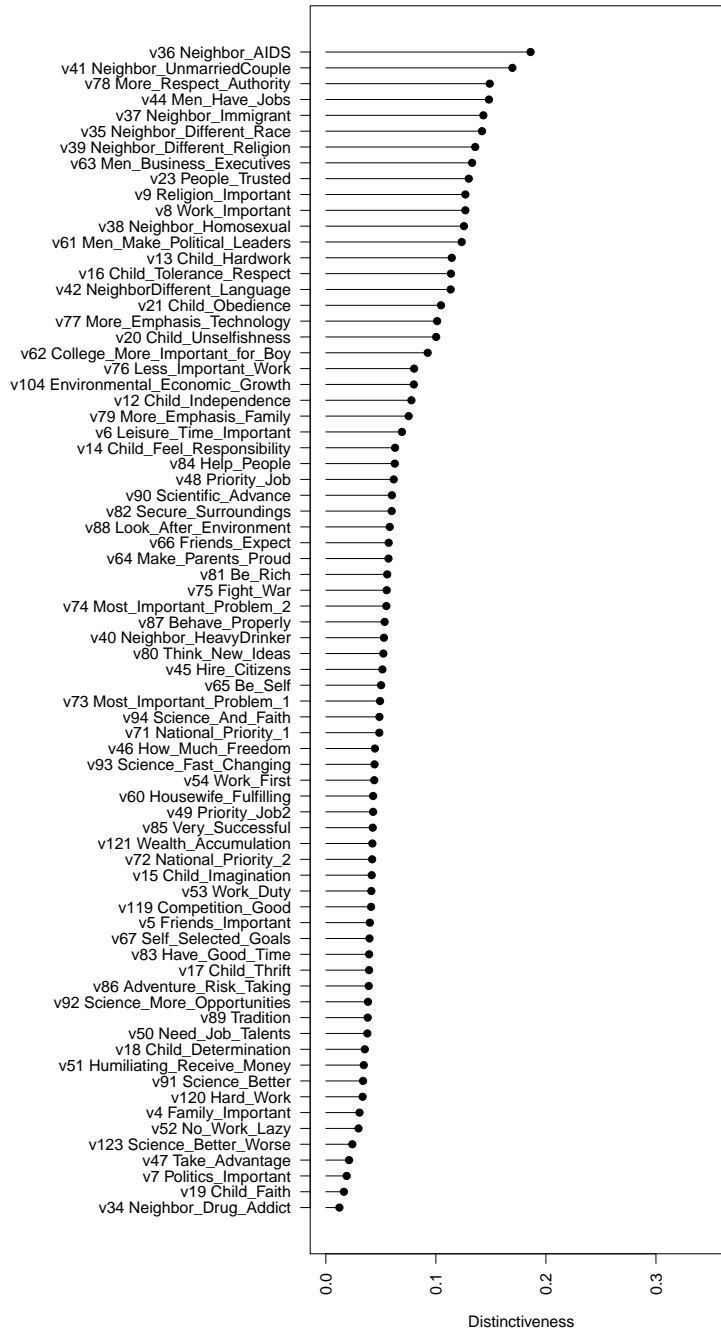
where we divide by $Q_r - 1$ because the final category is redundant.

Using this measure of distinctiveness, the following seven plots provide the distinctiveness for each question for each culture. The bigger the gap, the more distinct the response is. This bolsters the claims made in the main text and provides further details on what makes the groups different.

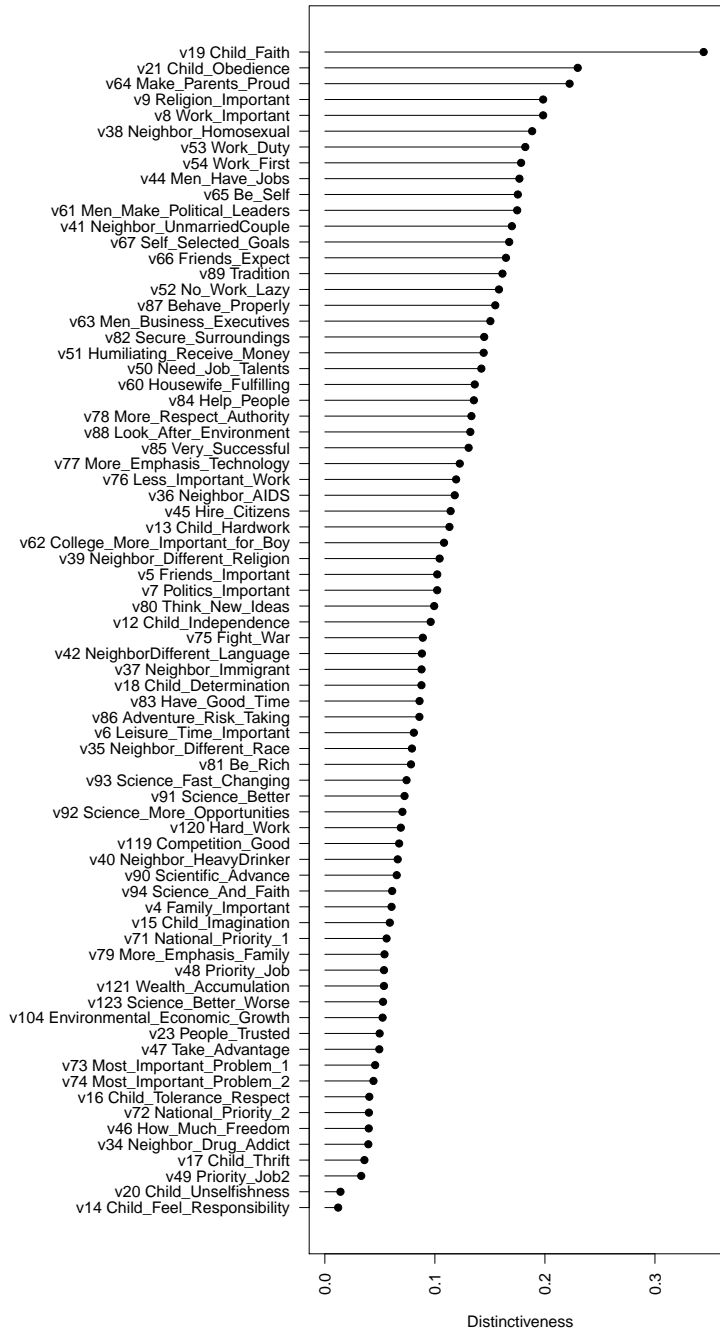
Group 1



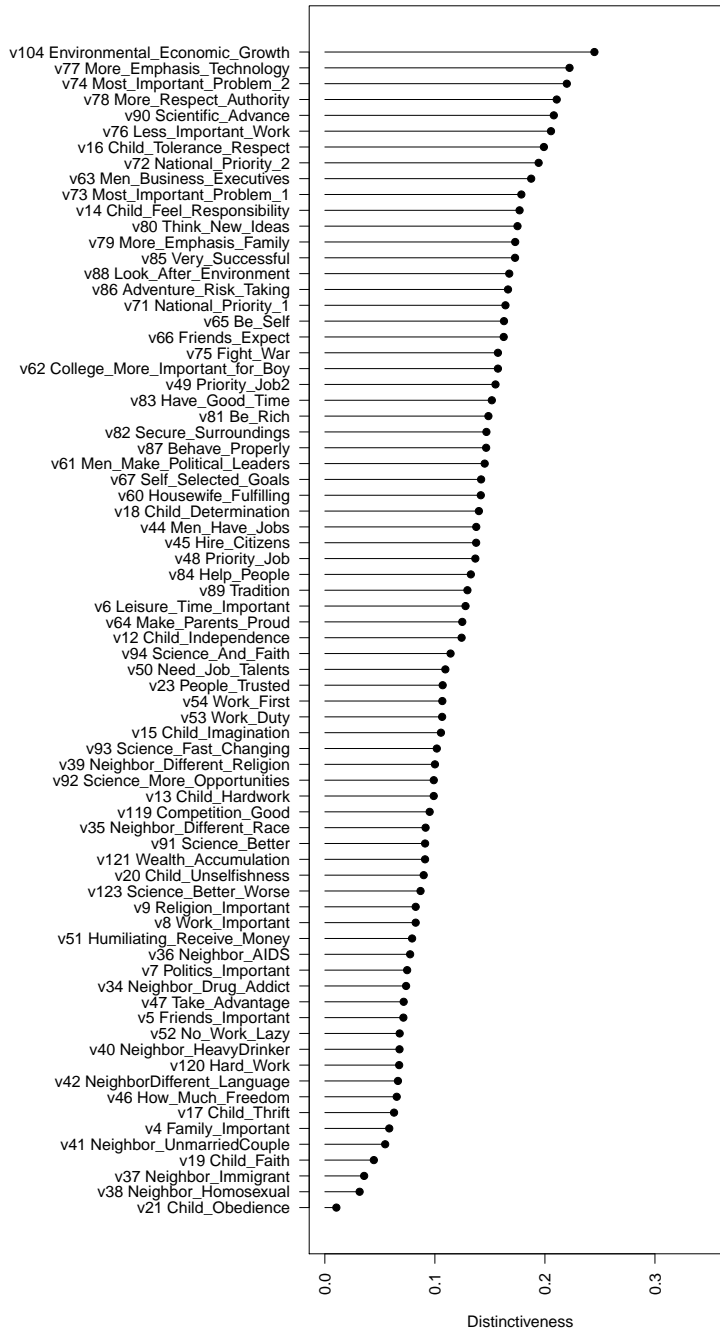
Group 2



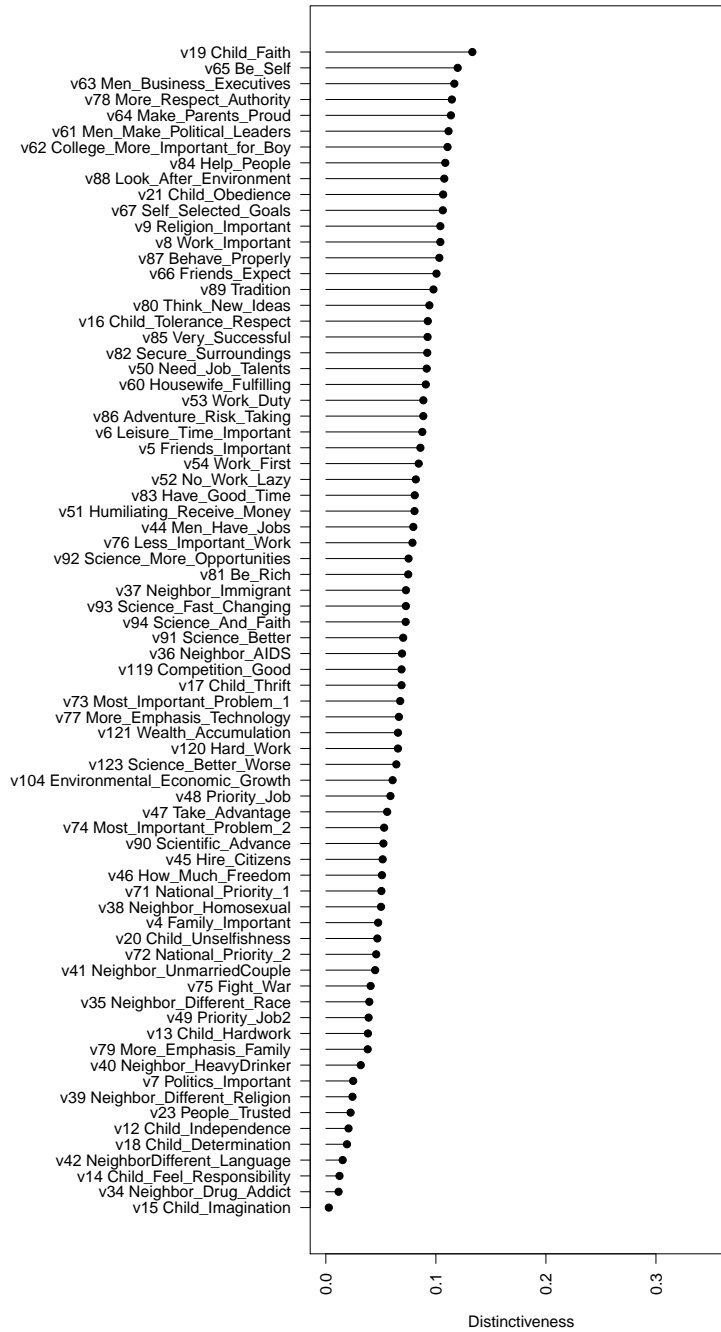
Group 3



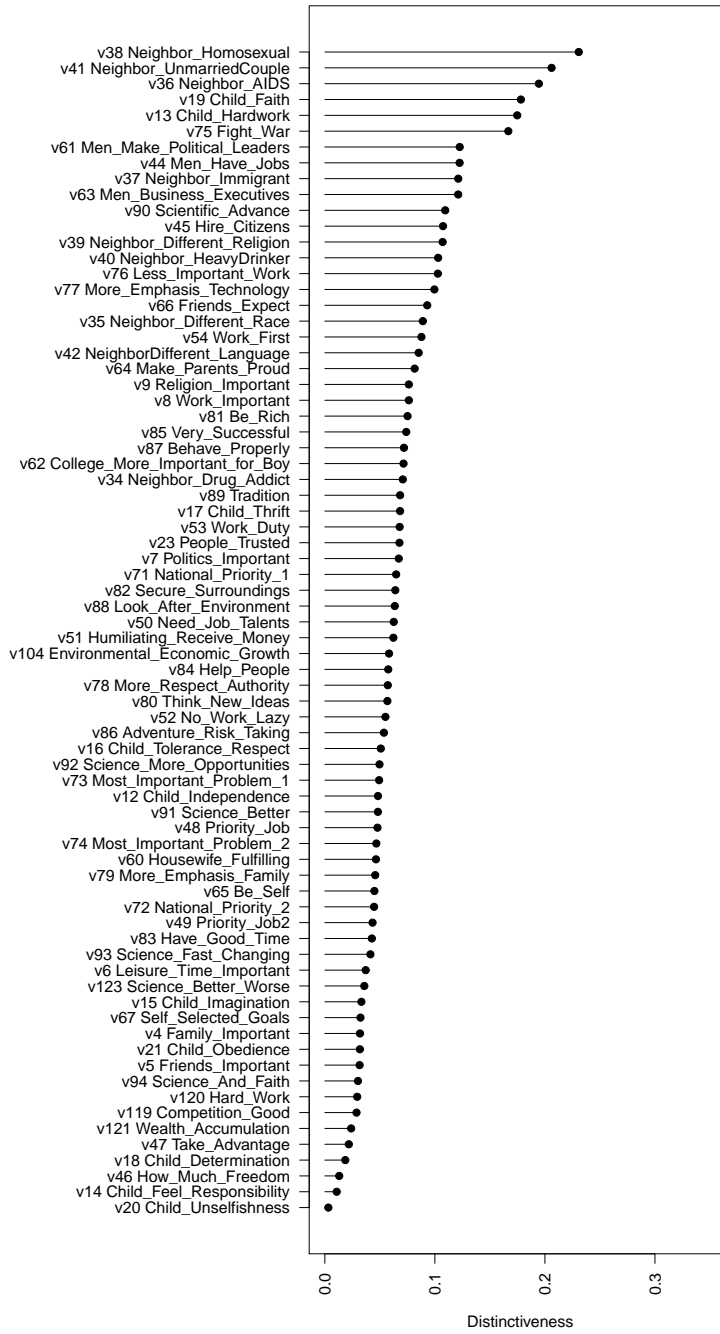
Group 4



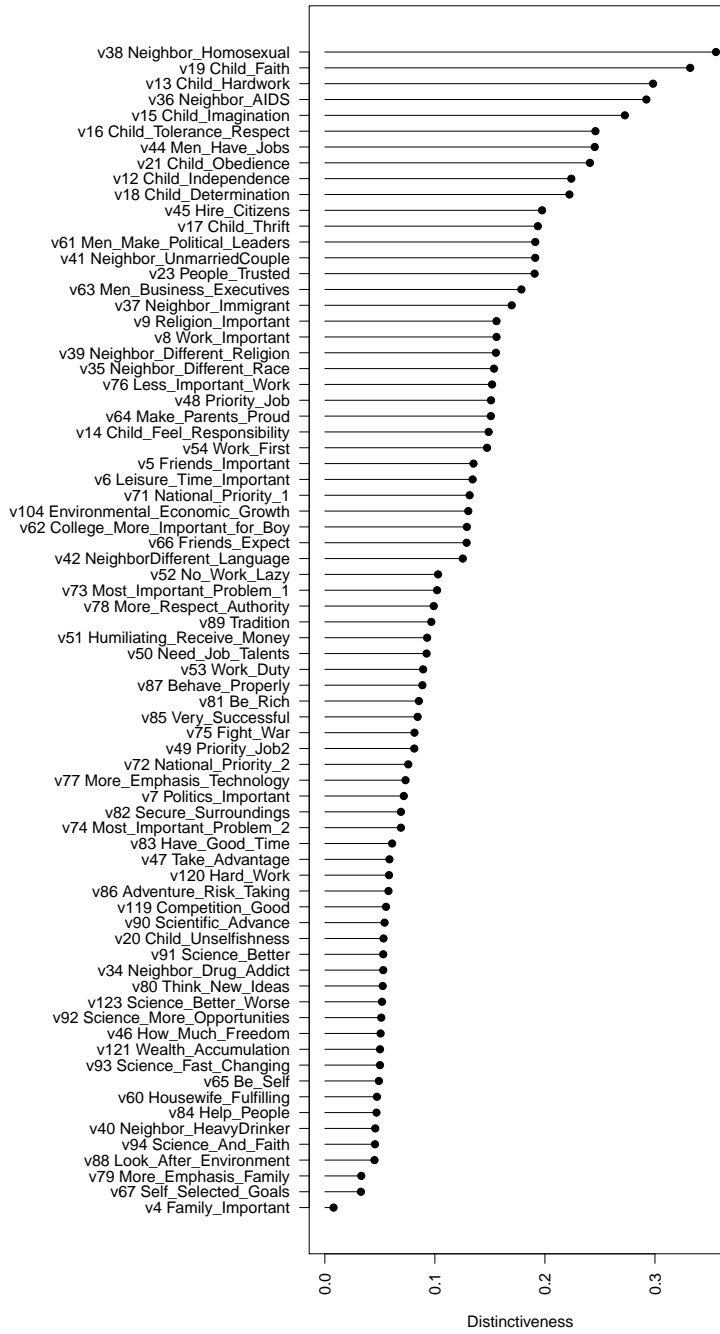
Group 5



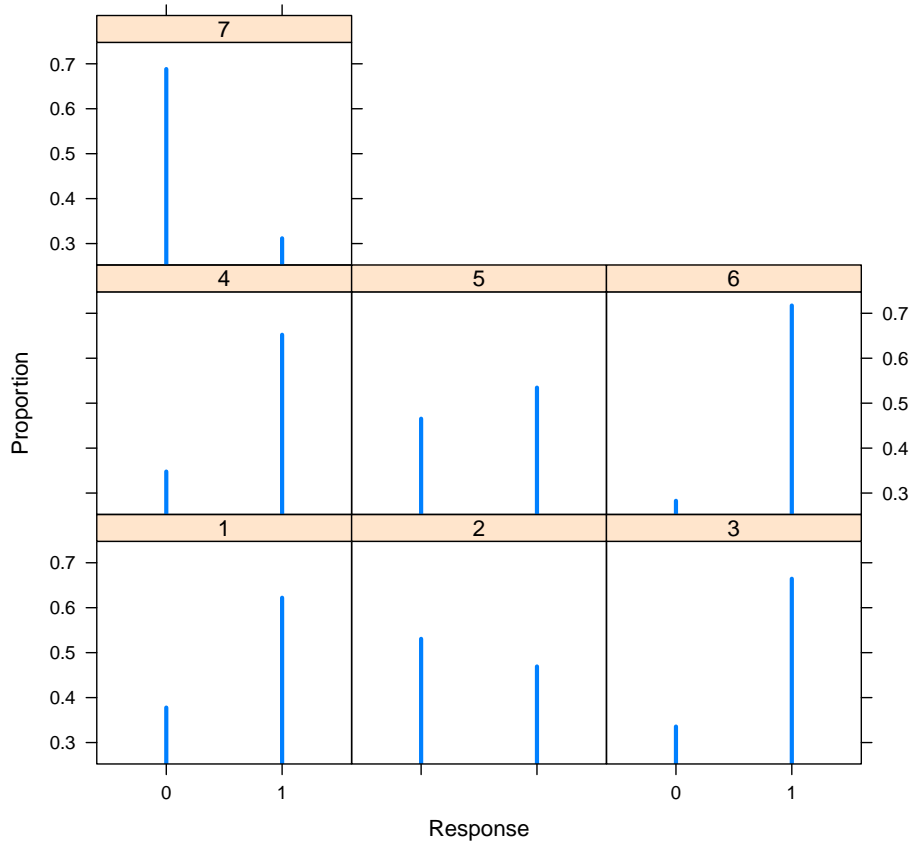
Group 6



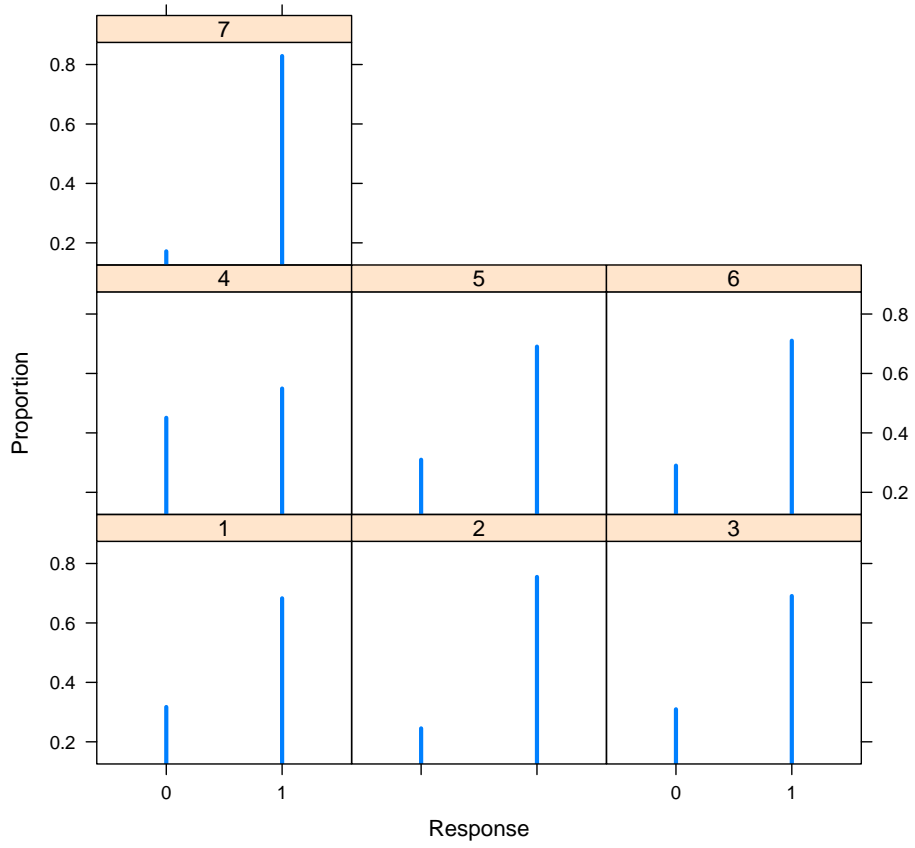
Group 7



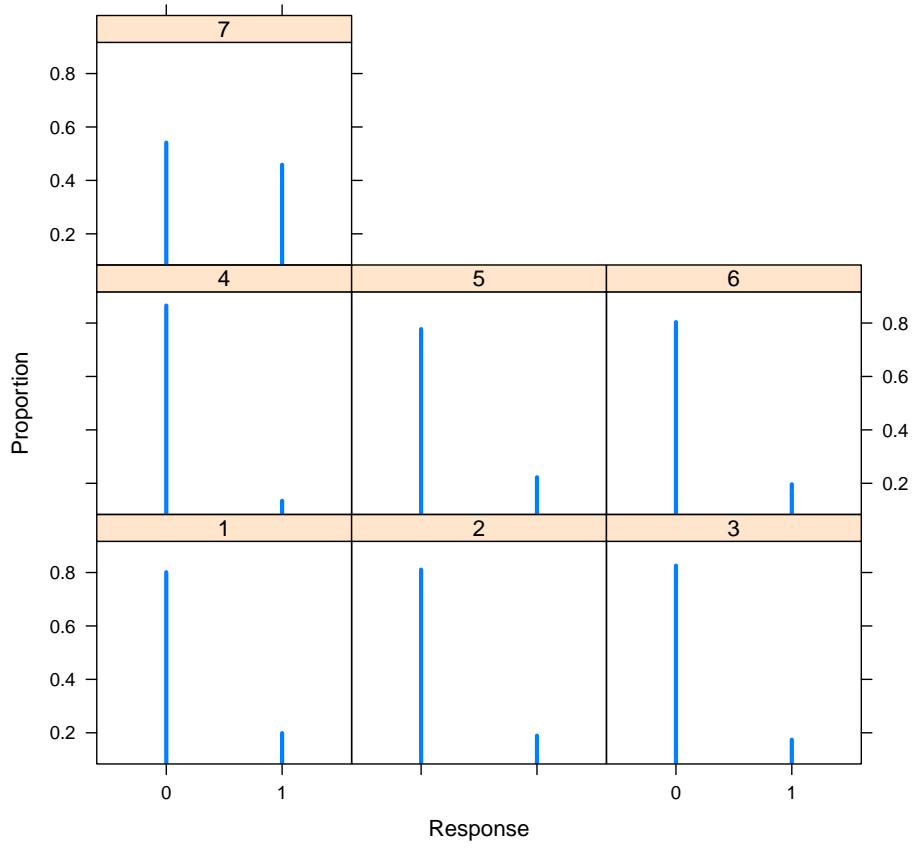
Child_Hardwork



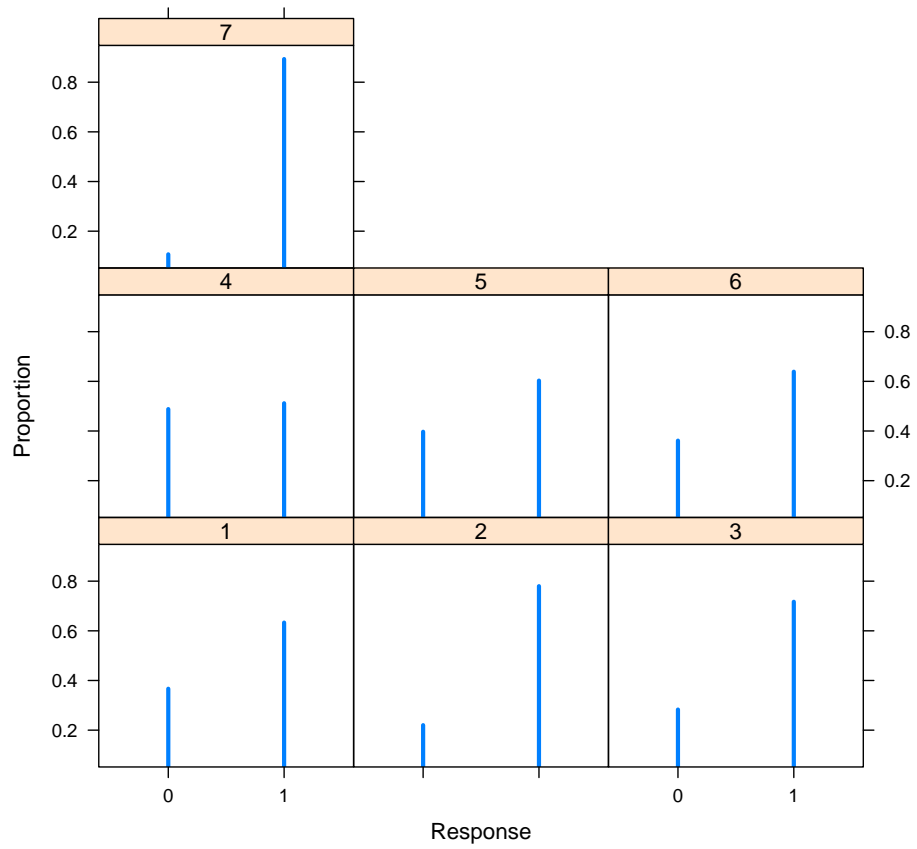
Child_Feel_Responsibility



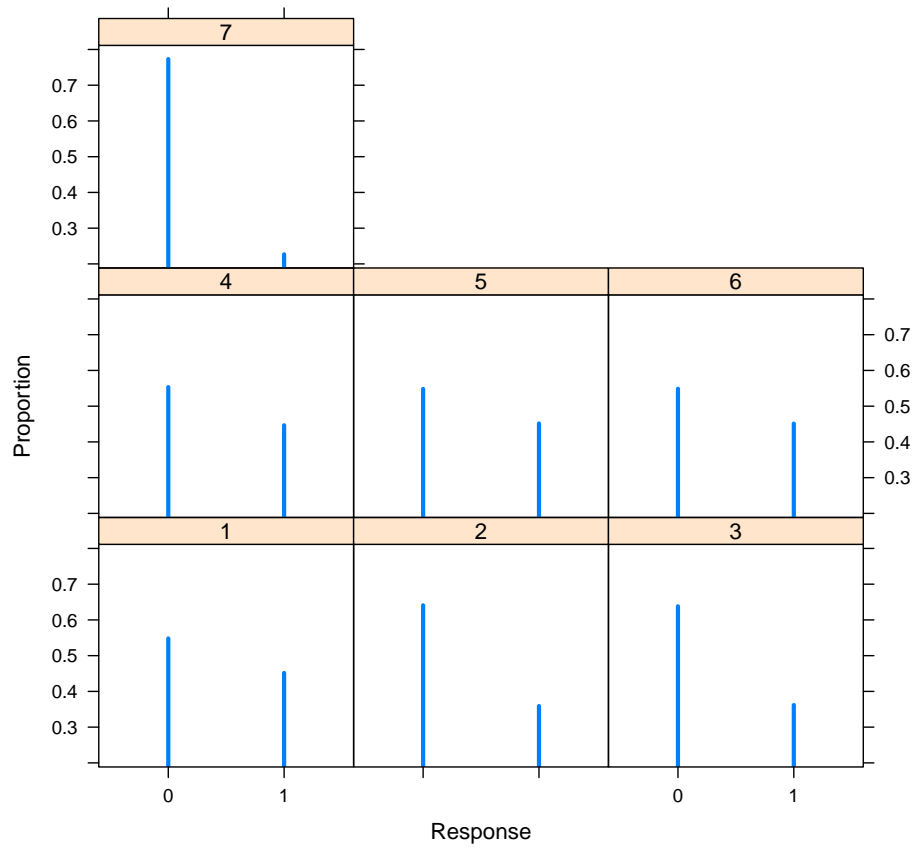
Child_Imagination



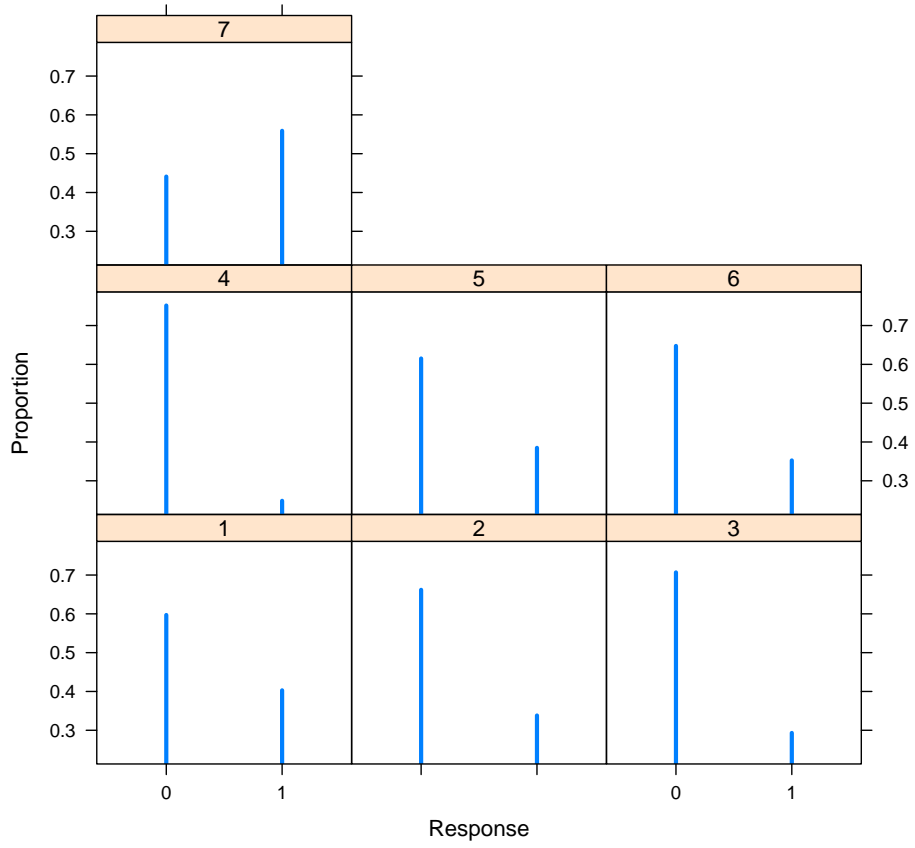
Child_Tolerance_Respect



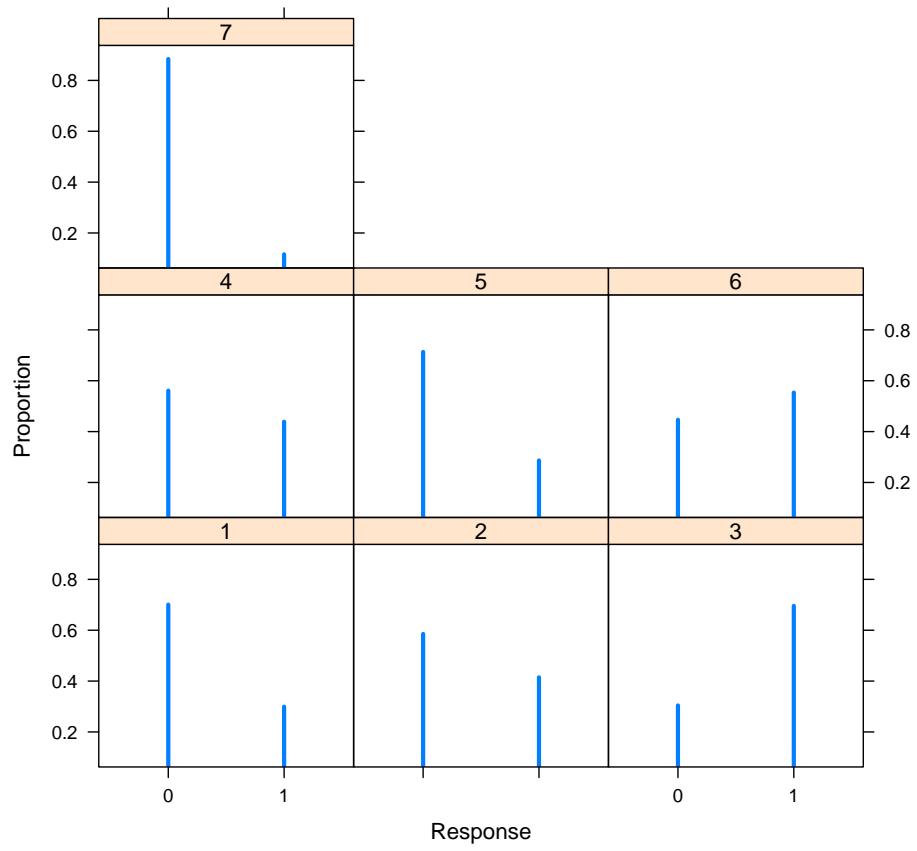
Child_Thrift



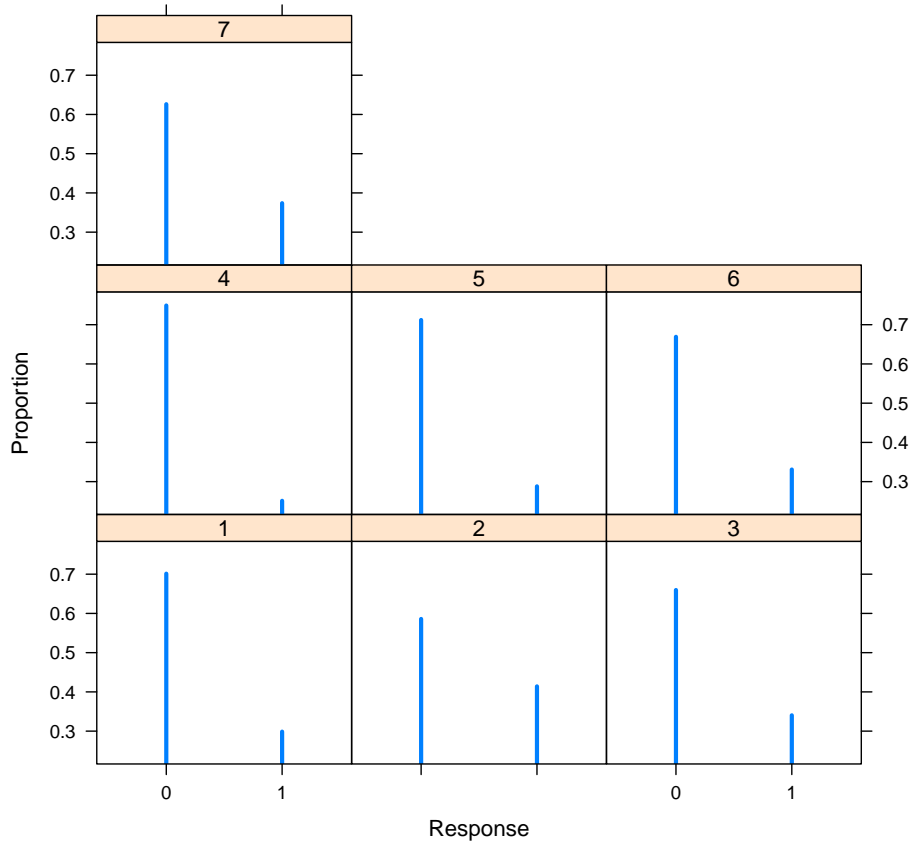
Child_Determination



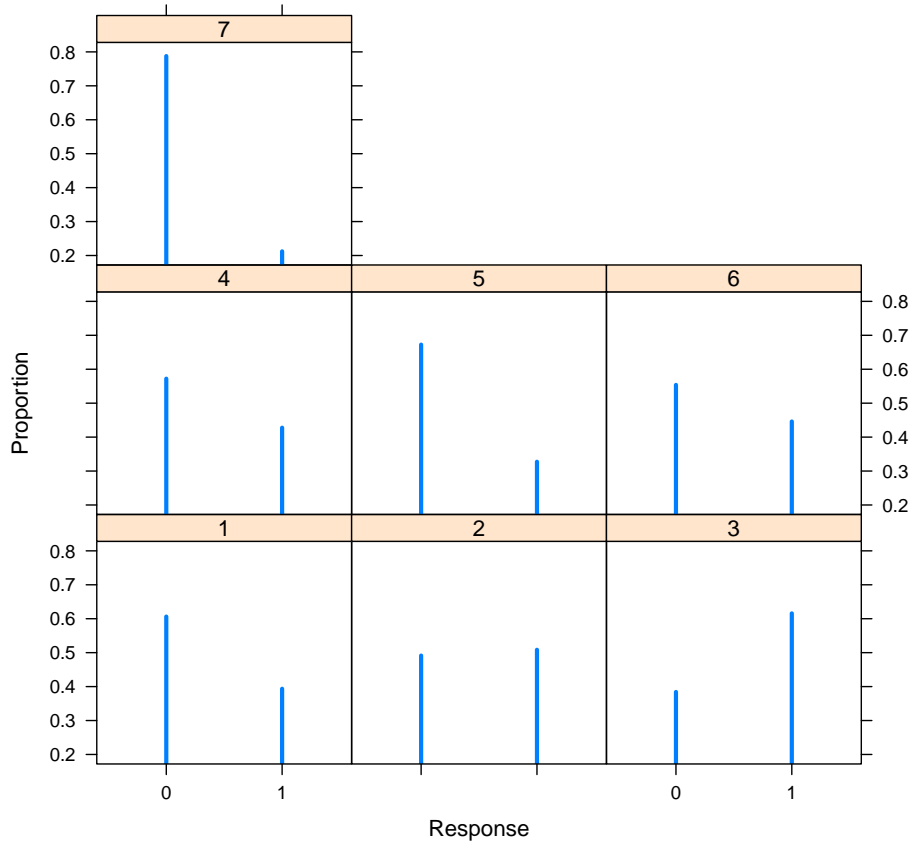
Child_Faith



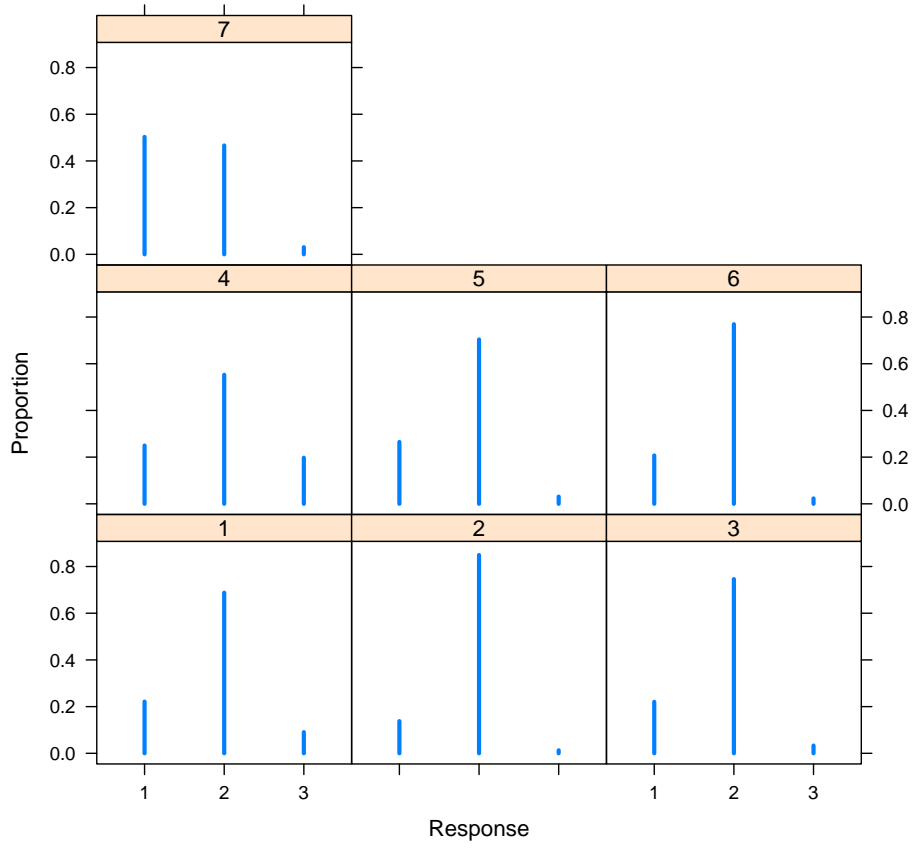
Child_Unselfishness



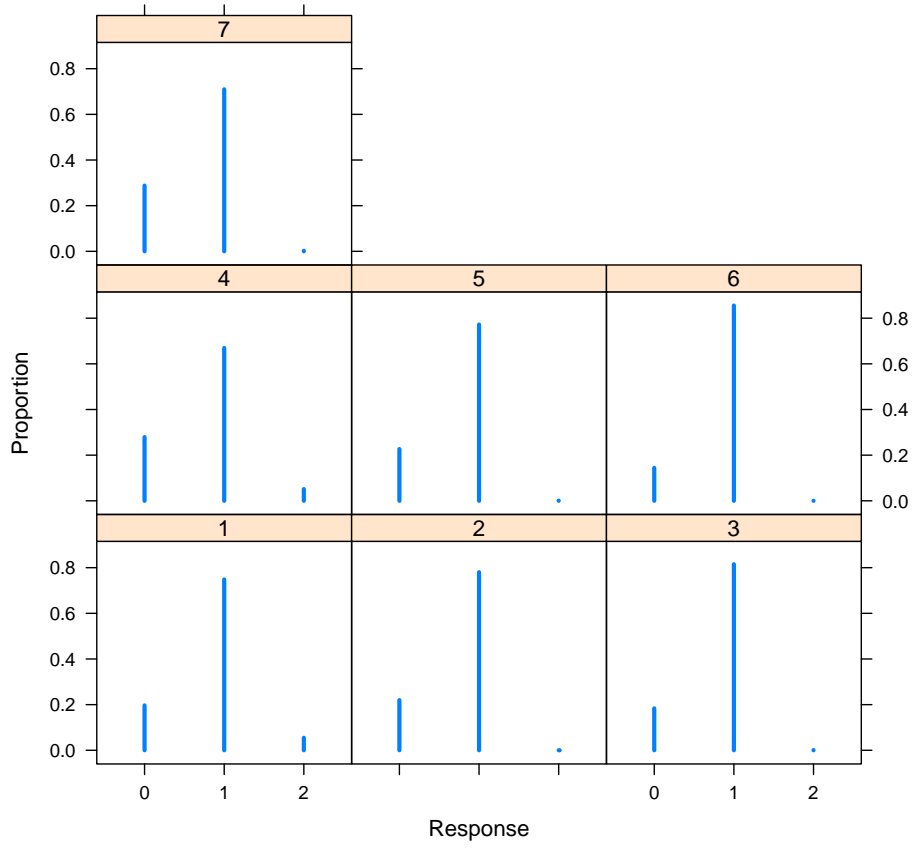
Child_Obedience



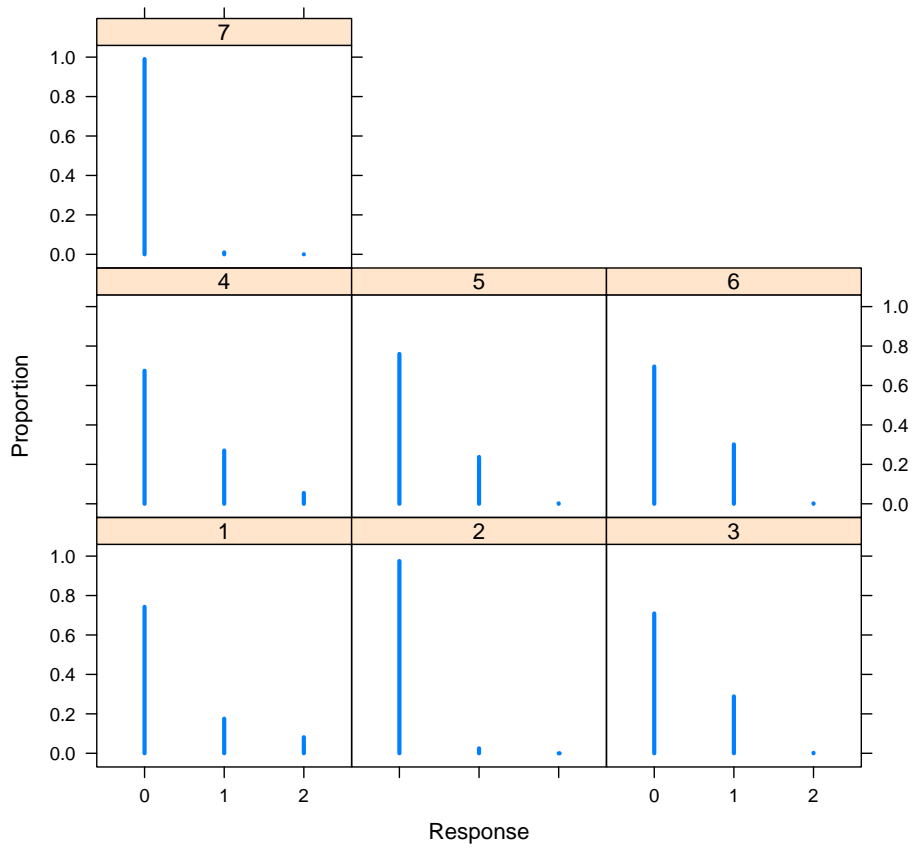
People_Trusted



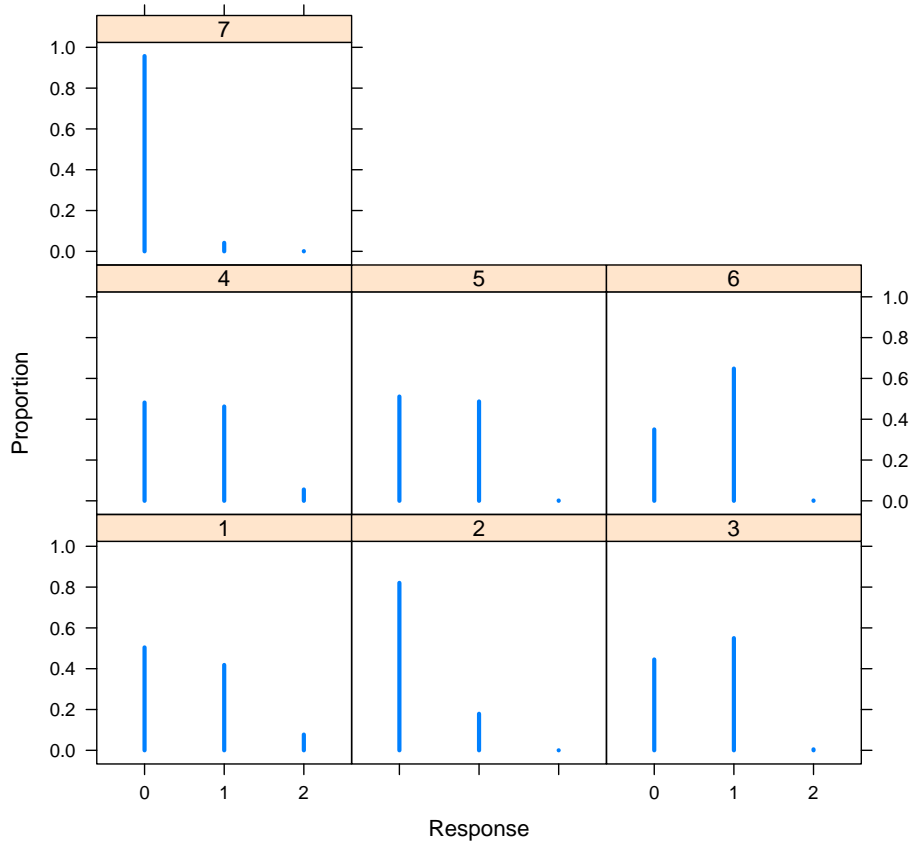
Neighbor_Drug_Addict



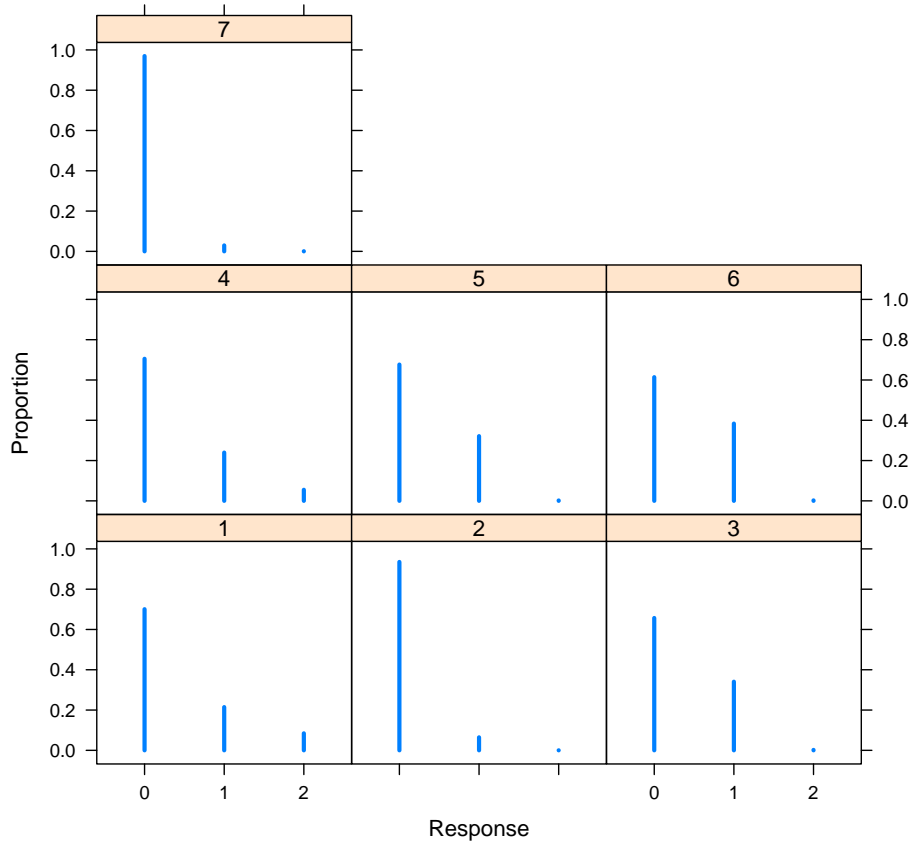
Neighbor_Different_Race



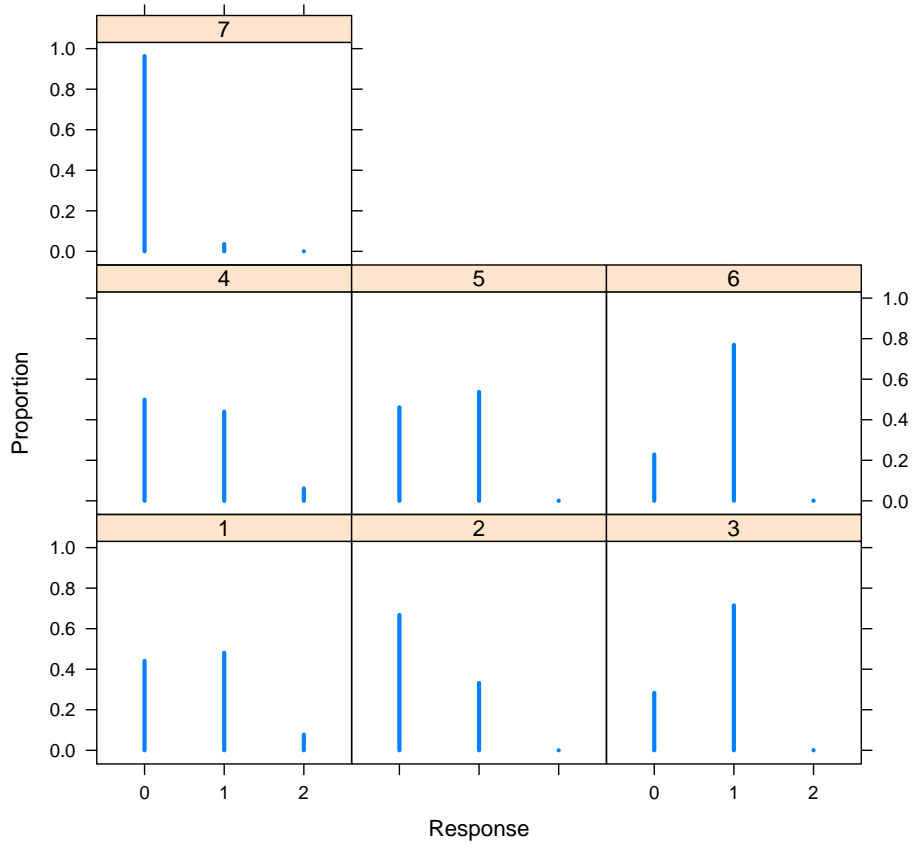
Neighbor_AIDS



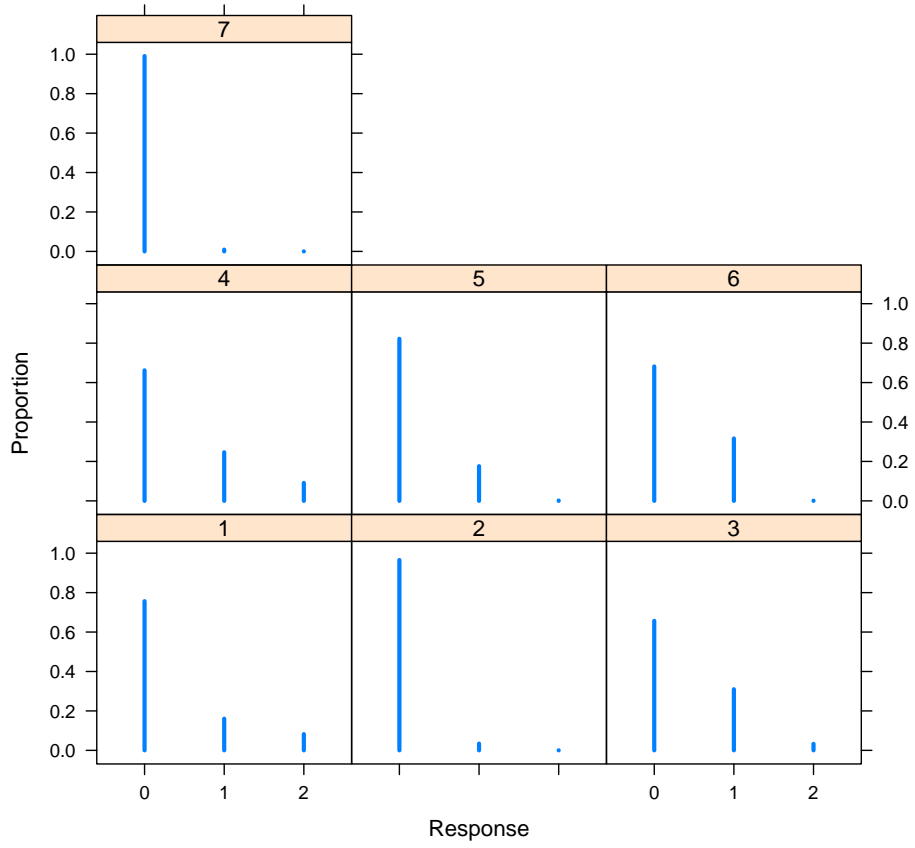
Neighbor_Immigrant



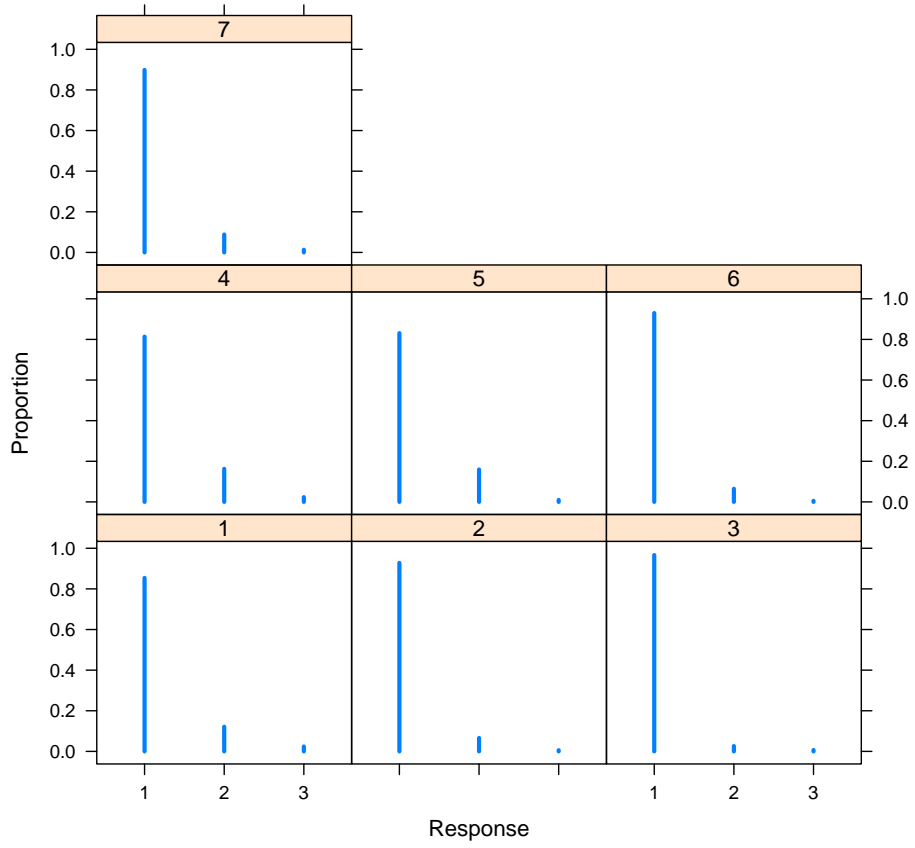
Neighbor_Homosexual



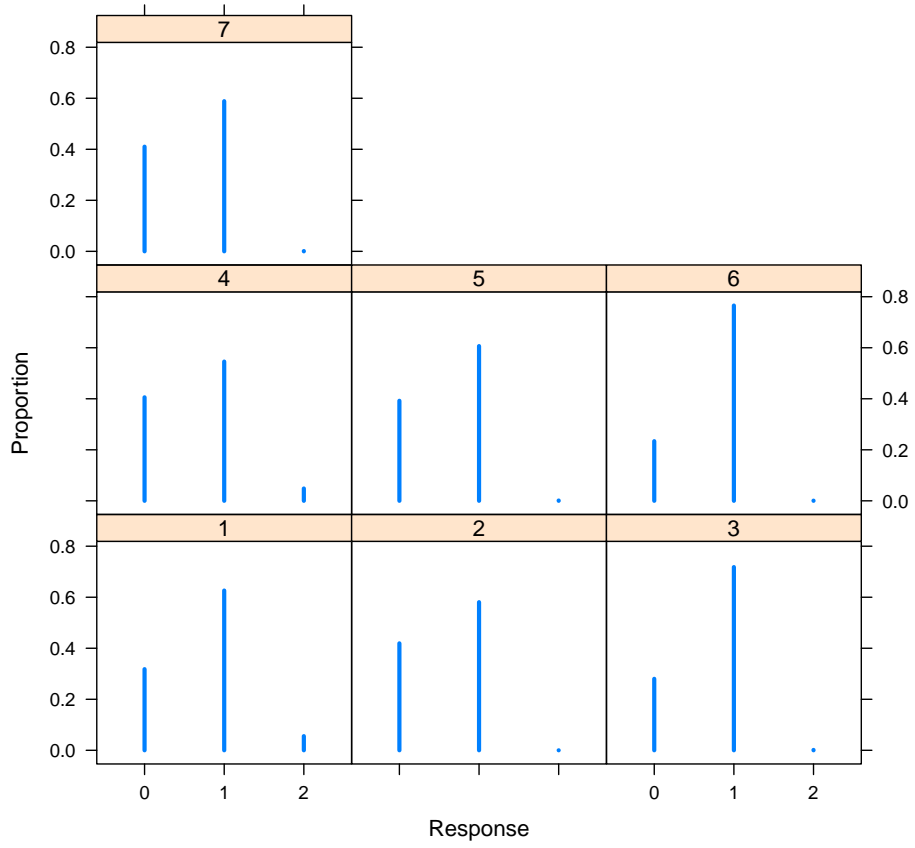
Neighbor_Different_Religion



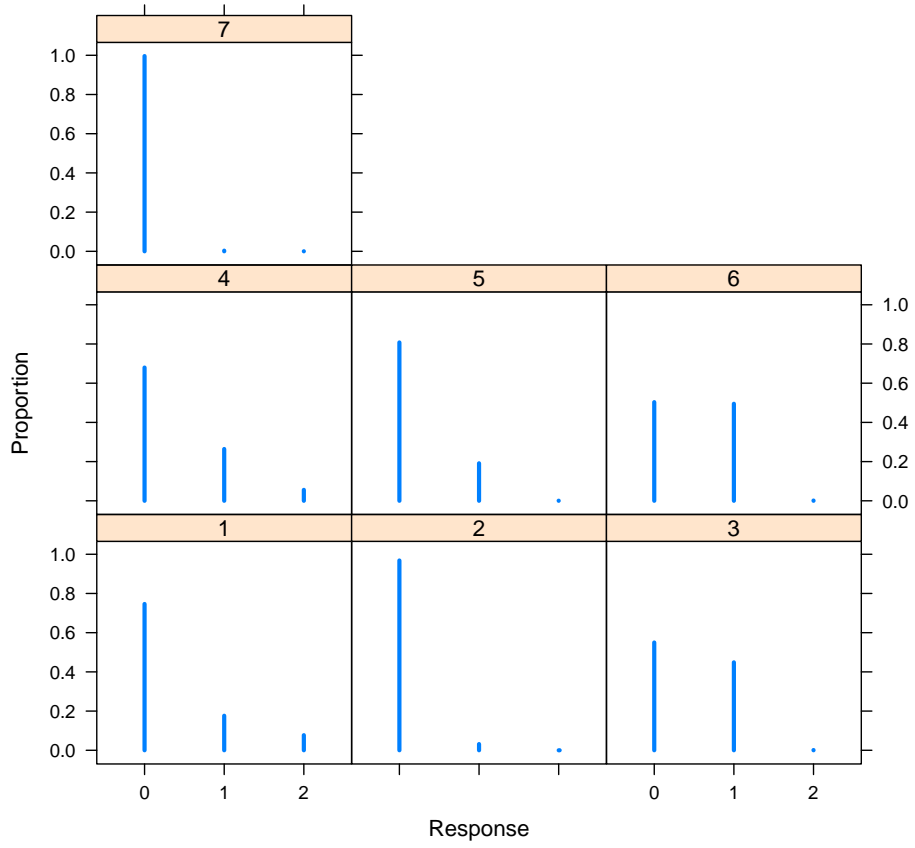
Family_Important



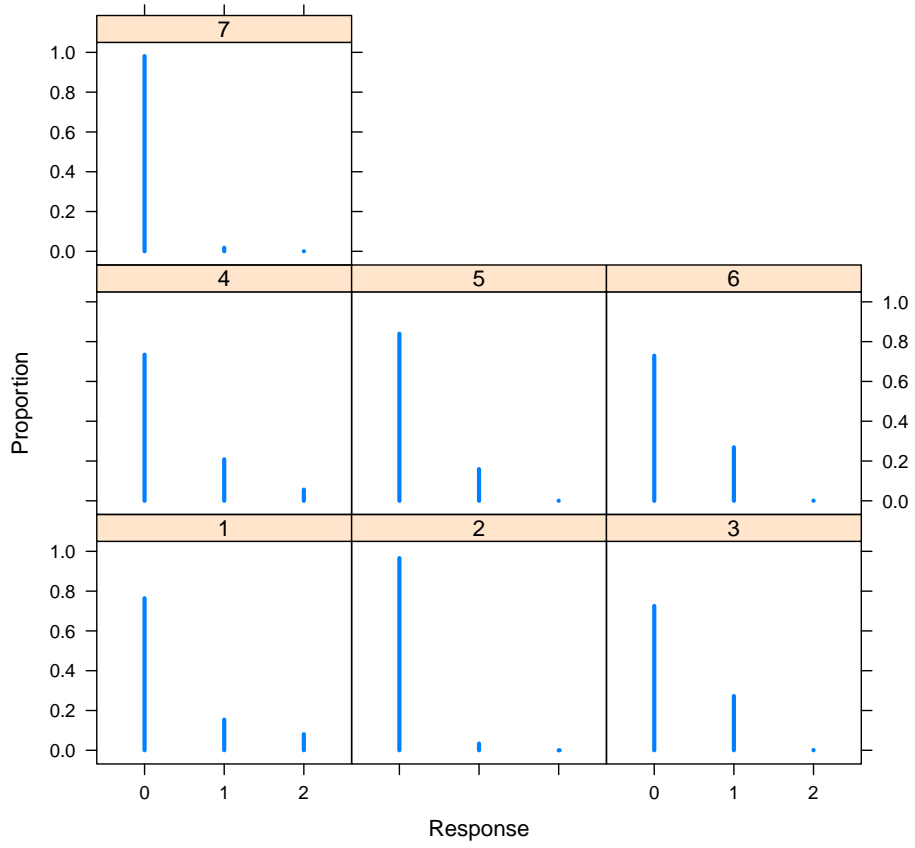
Neighbor_HeavyDrinker



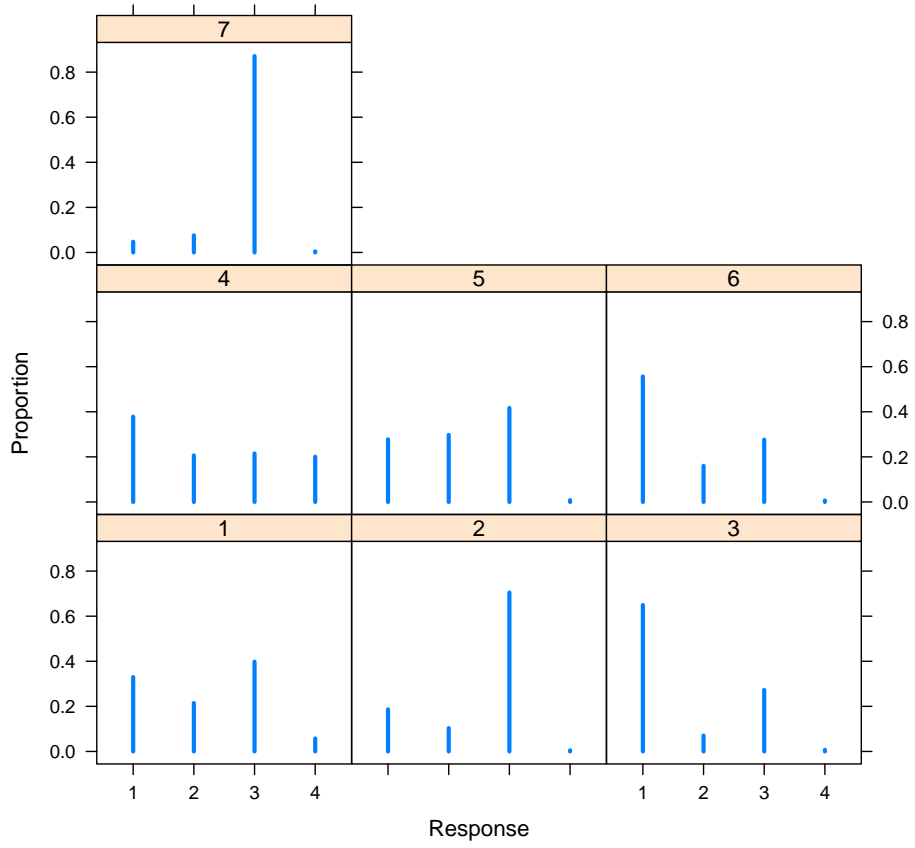
Neighbor_UnmarriedCouple



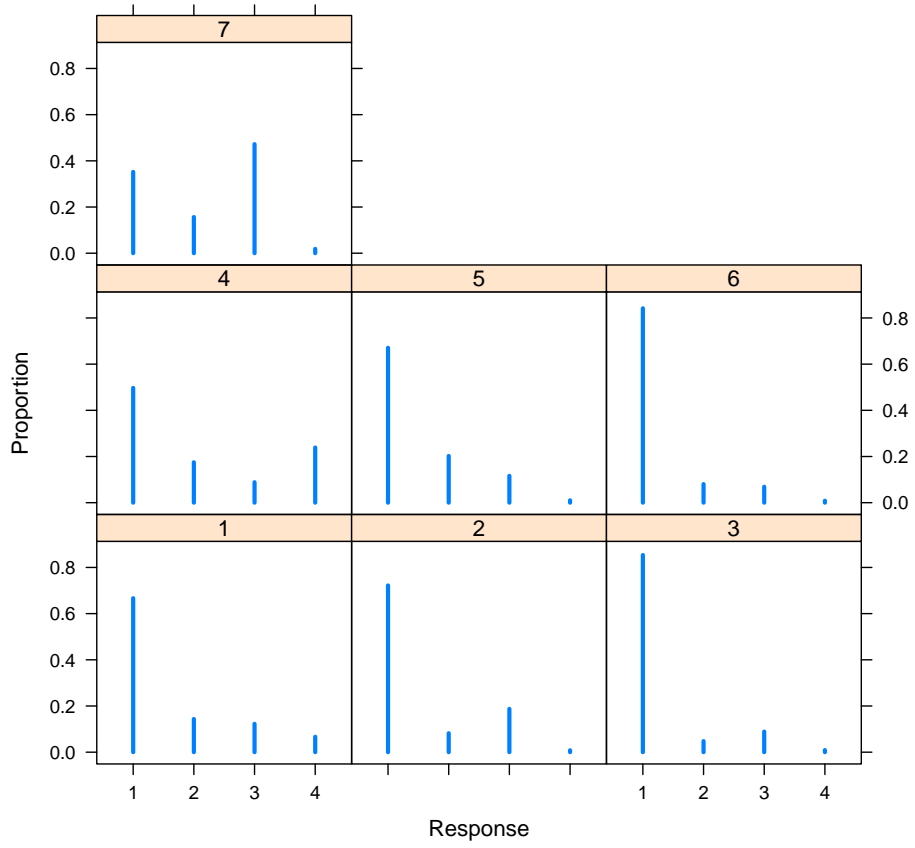
NeighborDifferent_Language



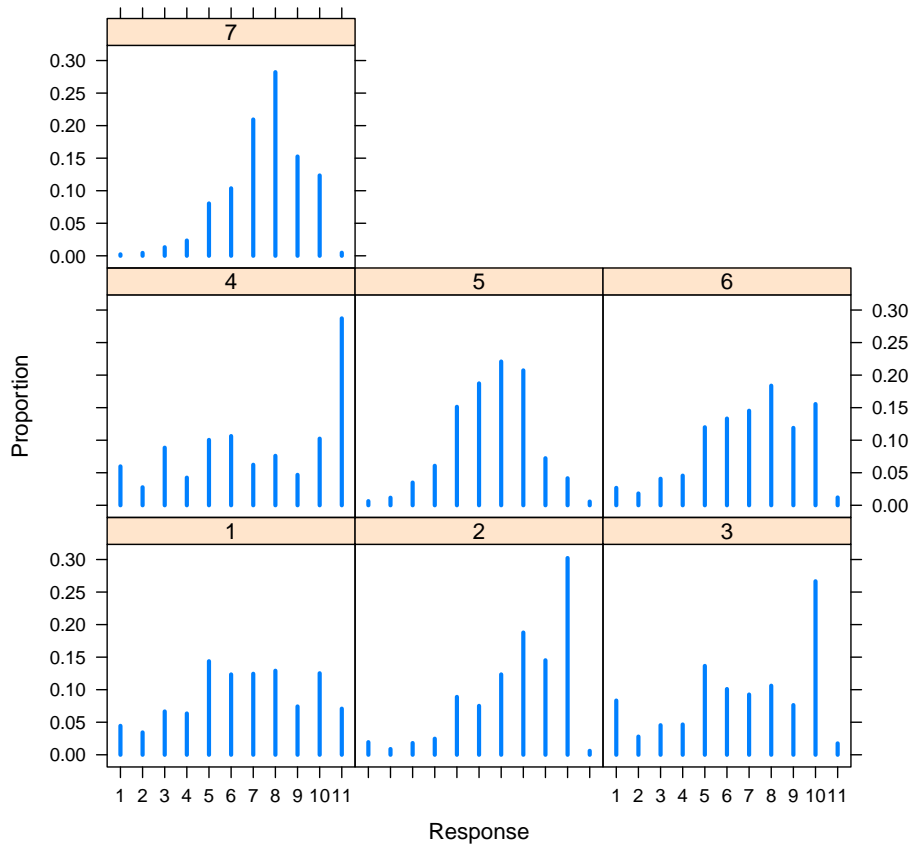
Men_Have_Jobs



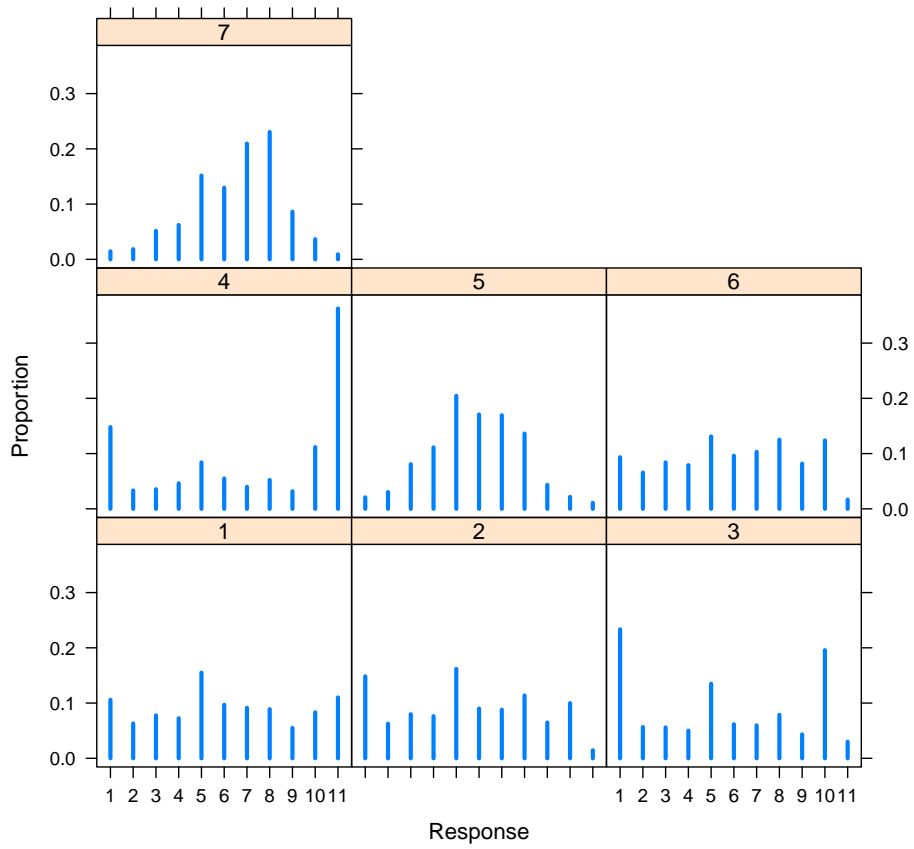
Hire_Citizens



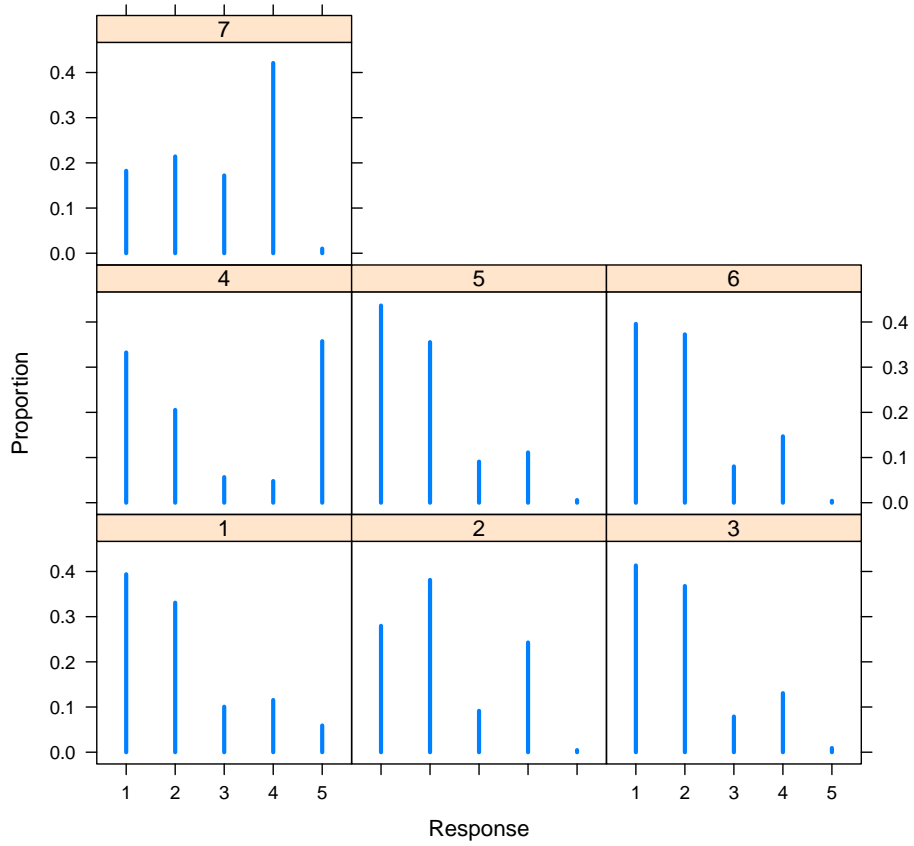
How_Much_Freedom



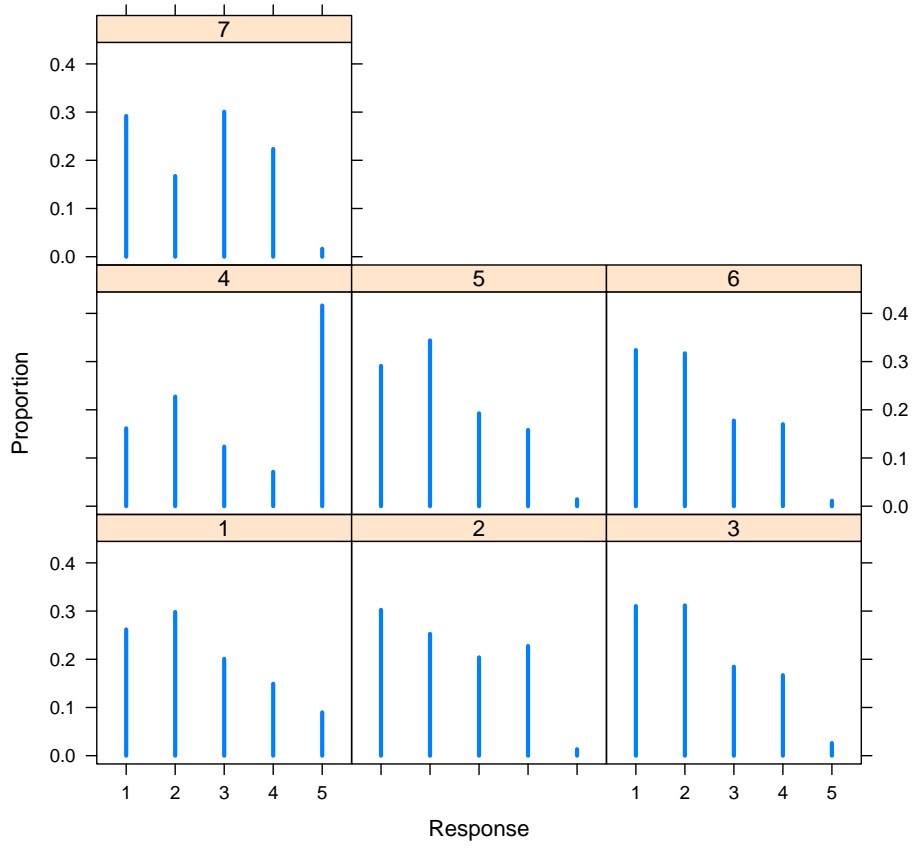
Take_Advantage



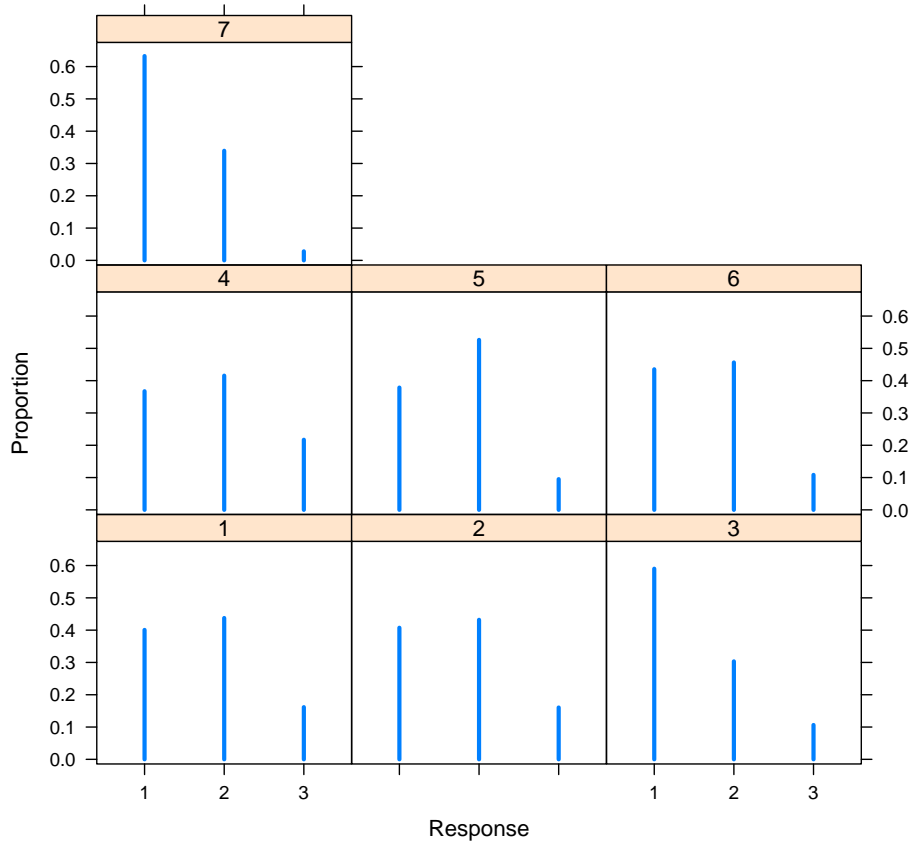
Priority_Job



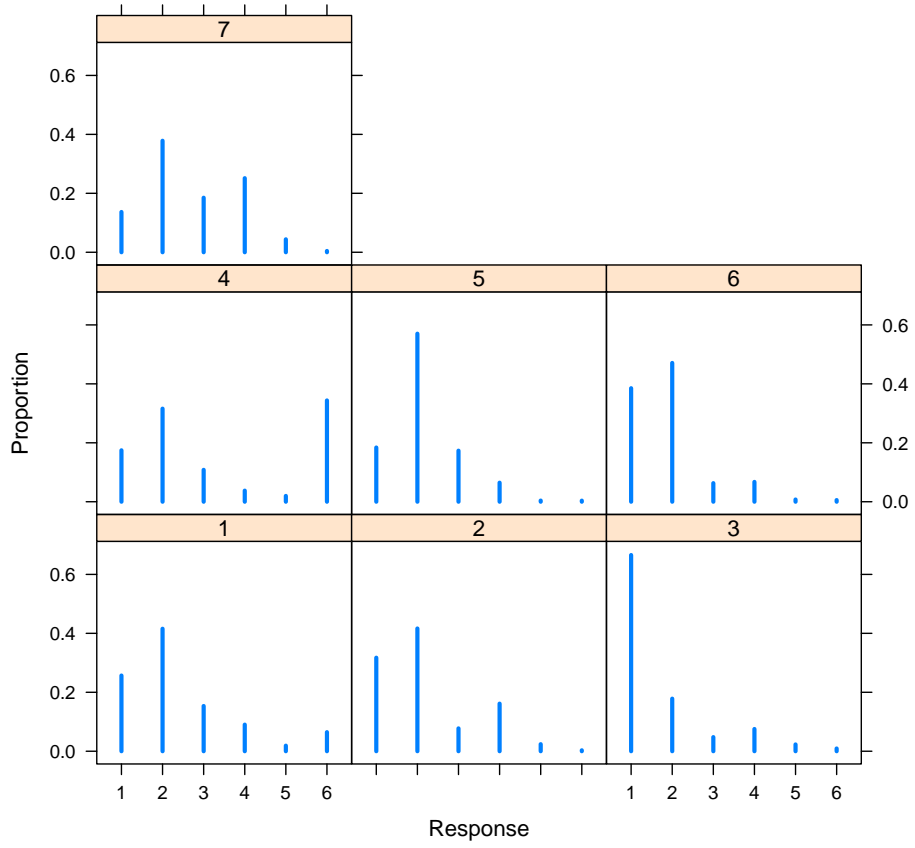
Priority_Job2



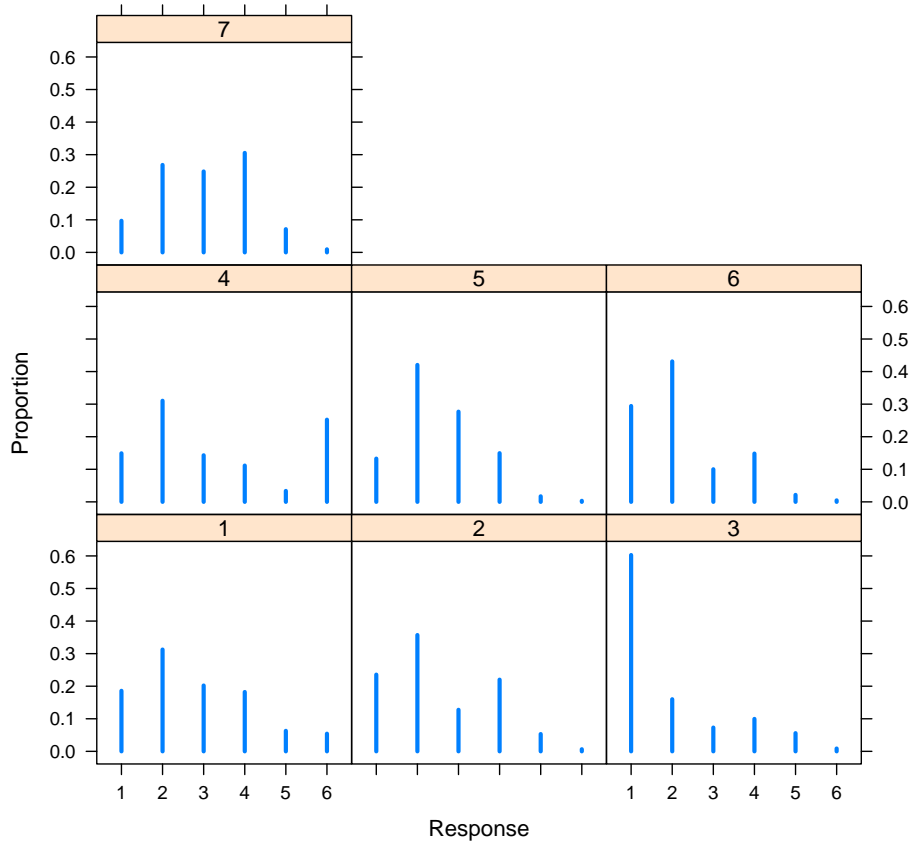
Friends_Important



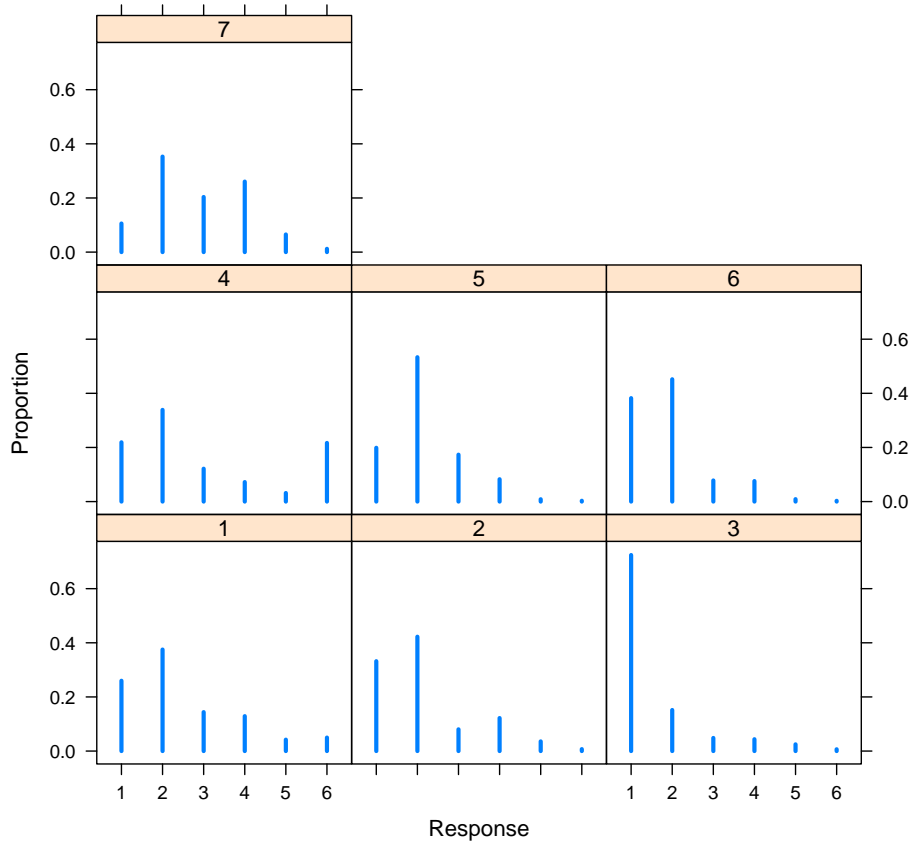
Need_Job_Talents



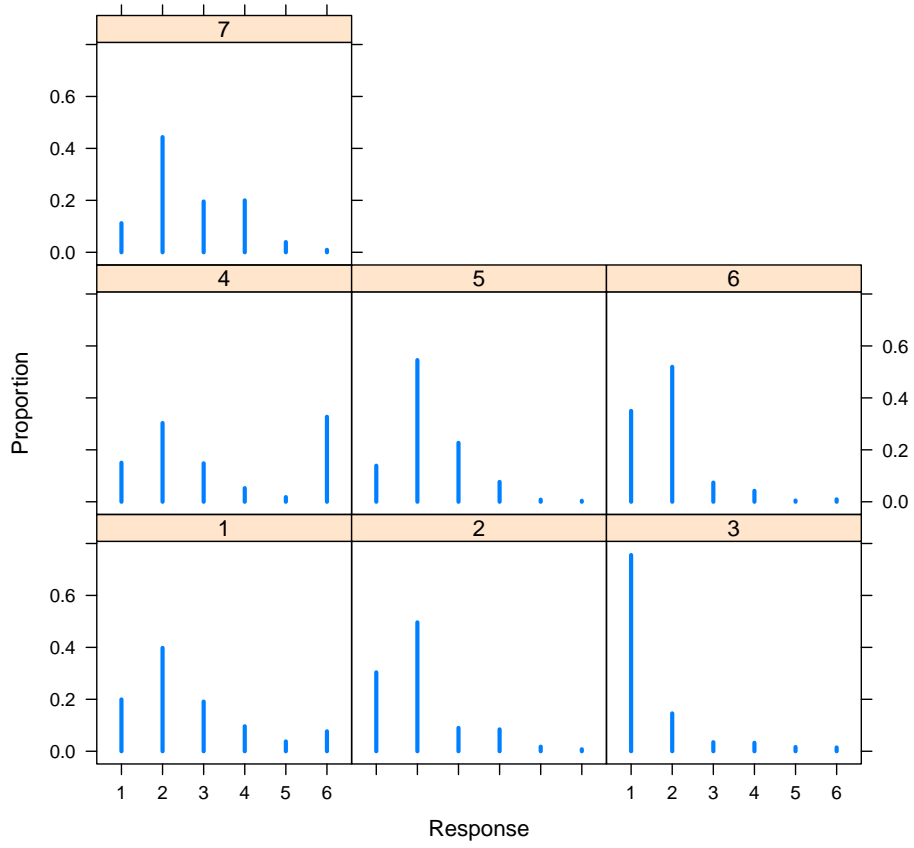
Humiliating_Receive_Money



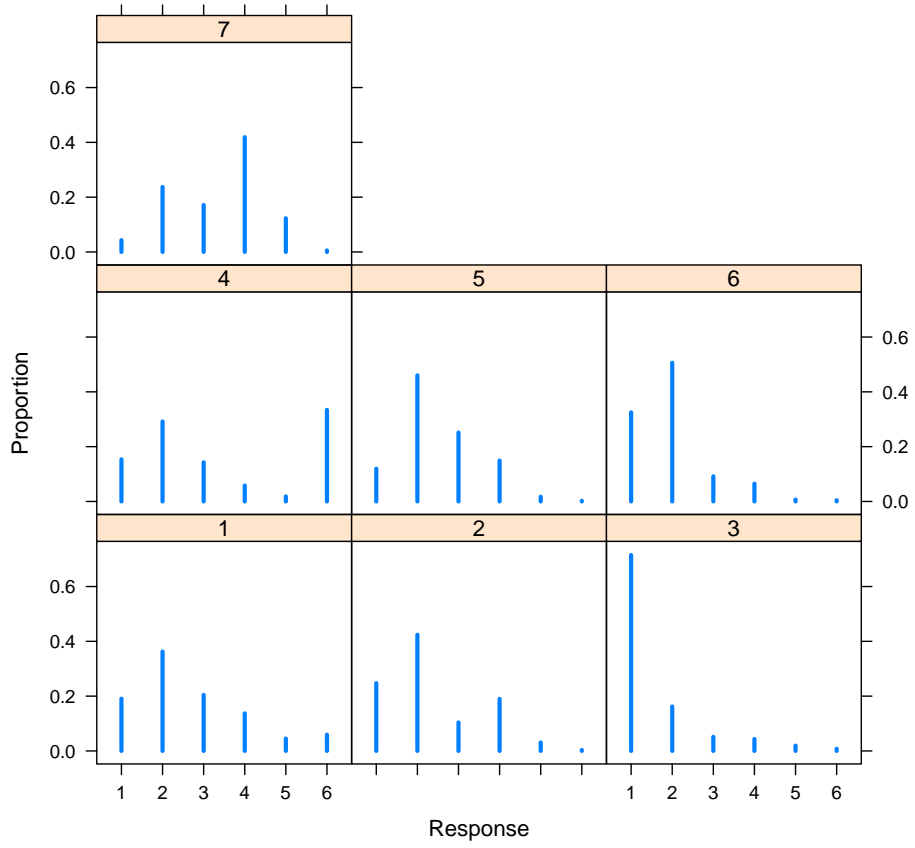
No_Work_Lazy



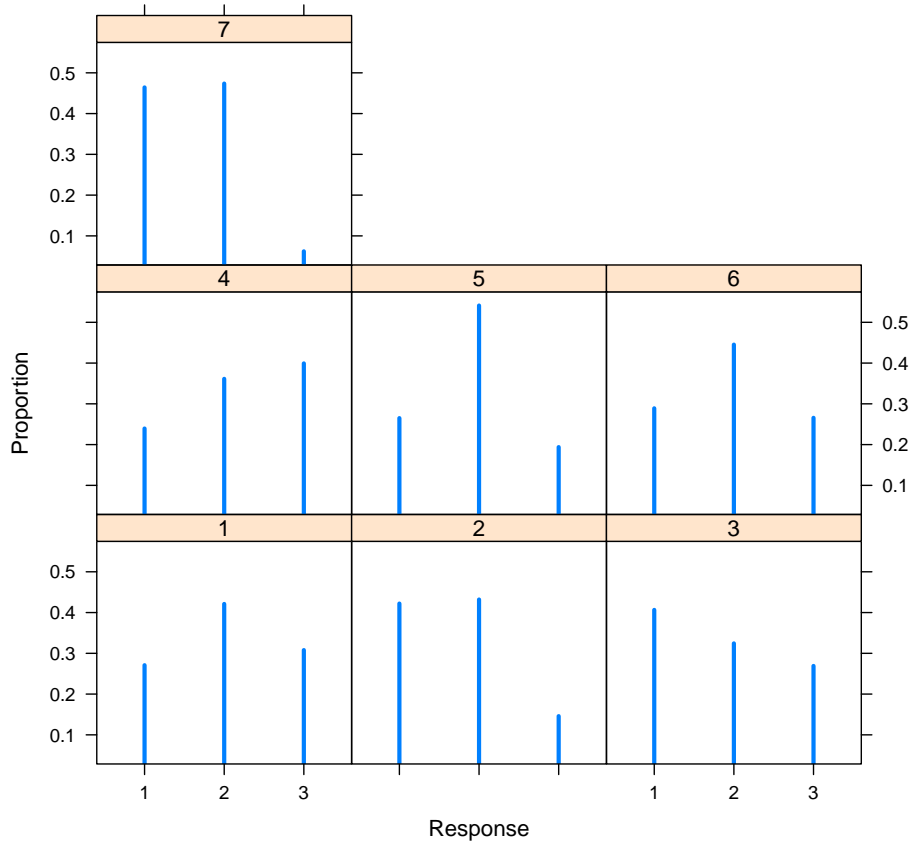
Work_Duty



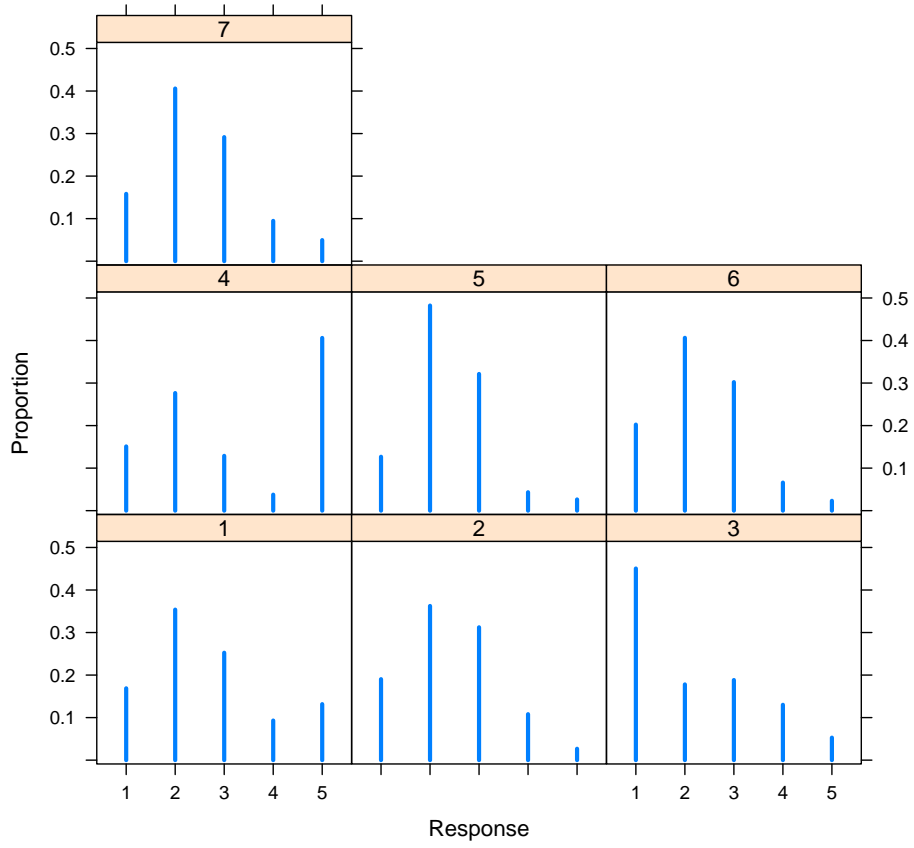
Work_First



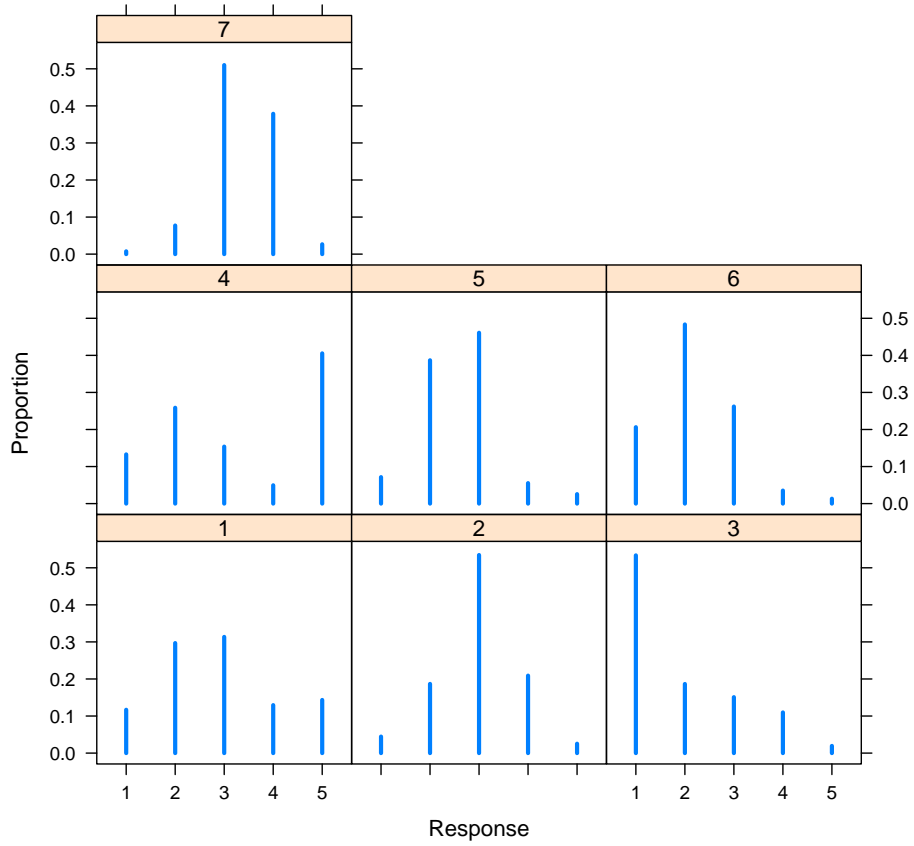
Leisure_Time_Important



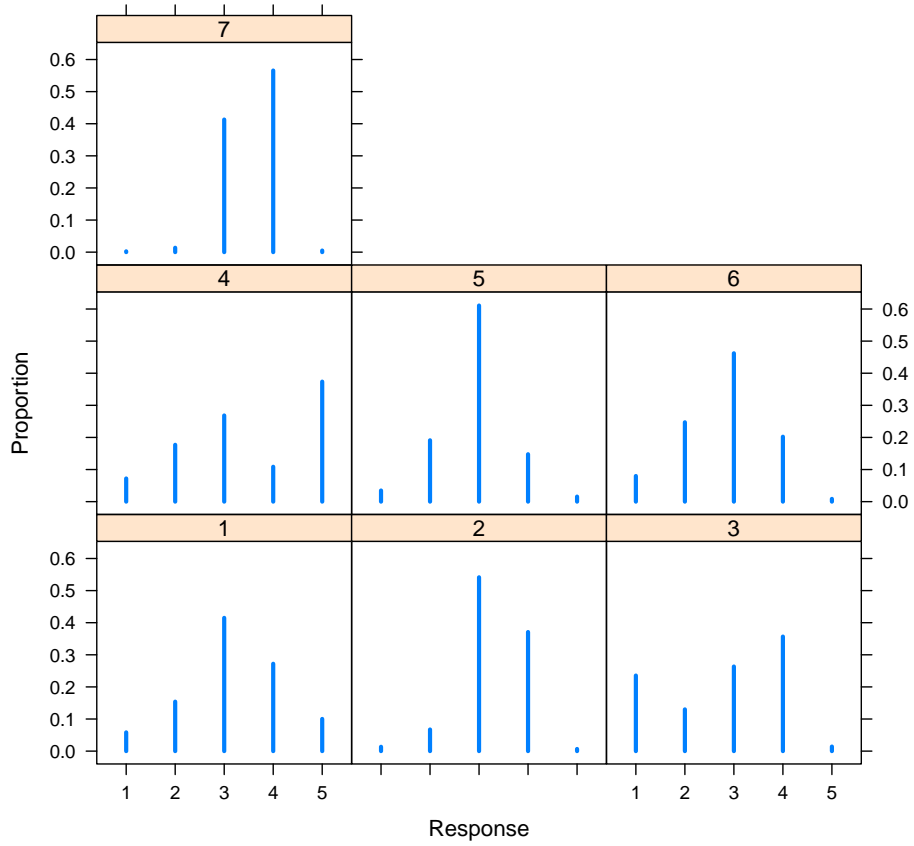
Housewife_Fulfilling



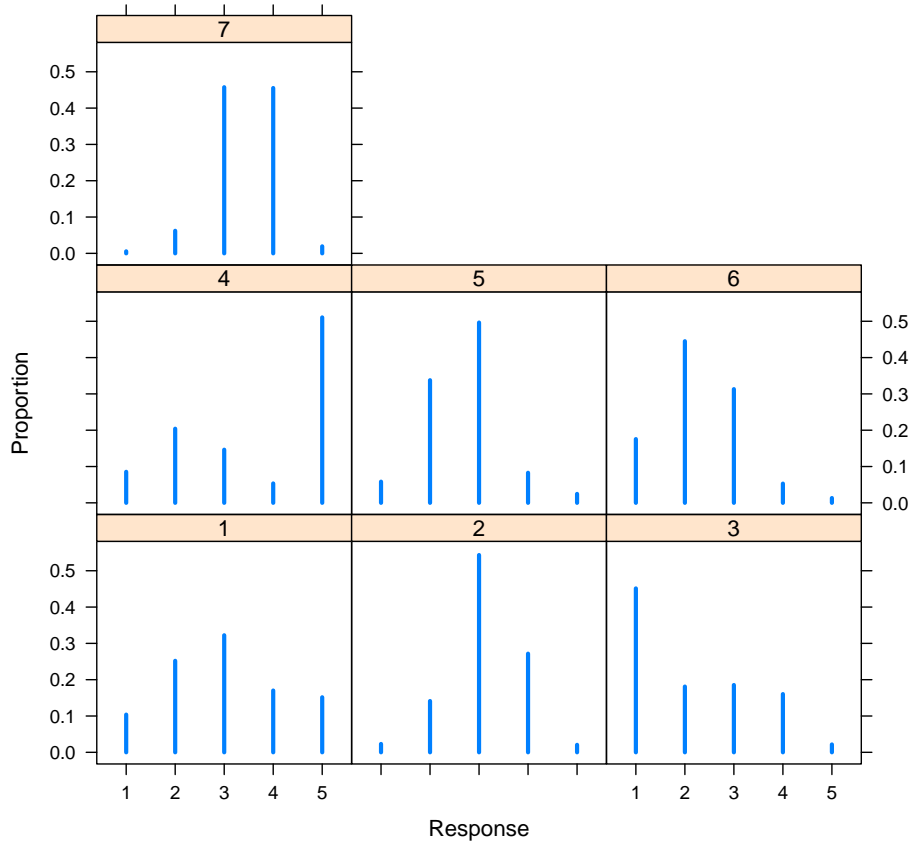
Men_Make_Political_Leaders



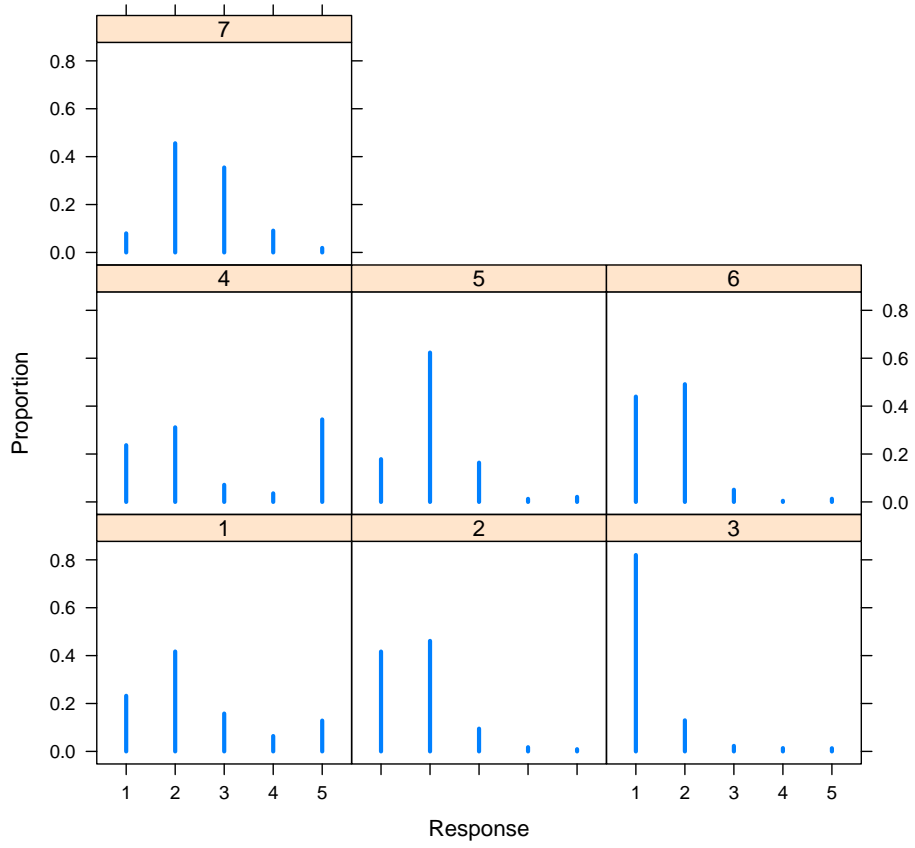
College_More_Important_for_Boy



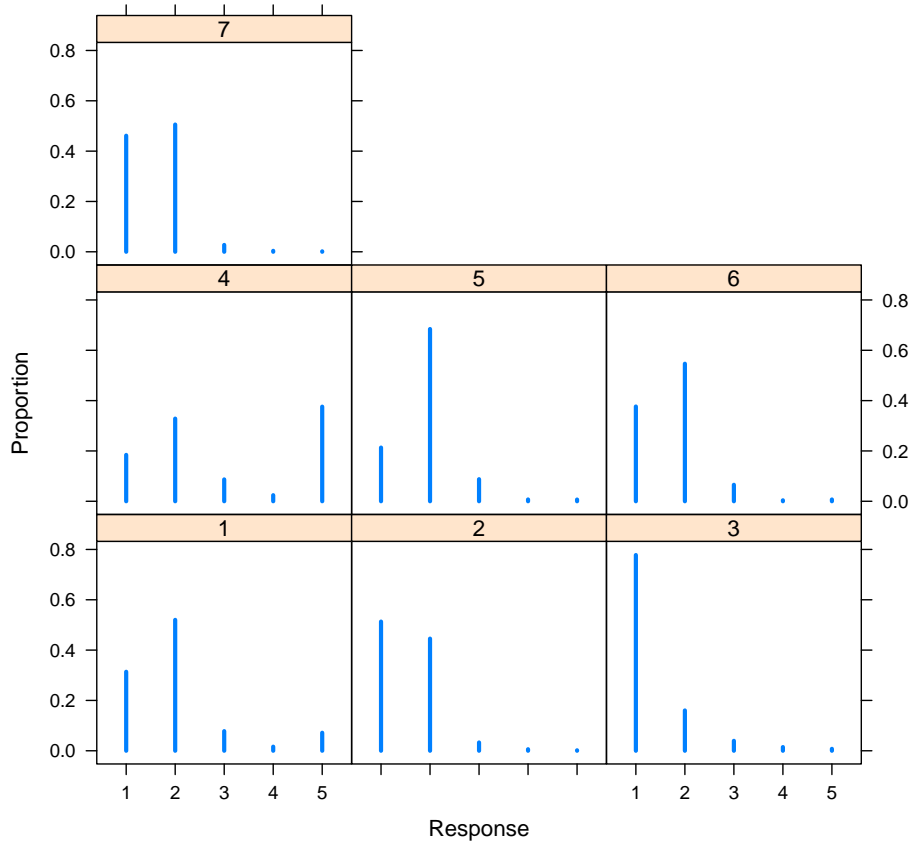
Men_Business_Executives



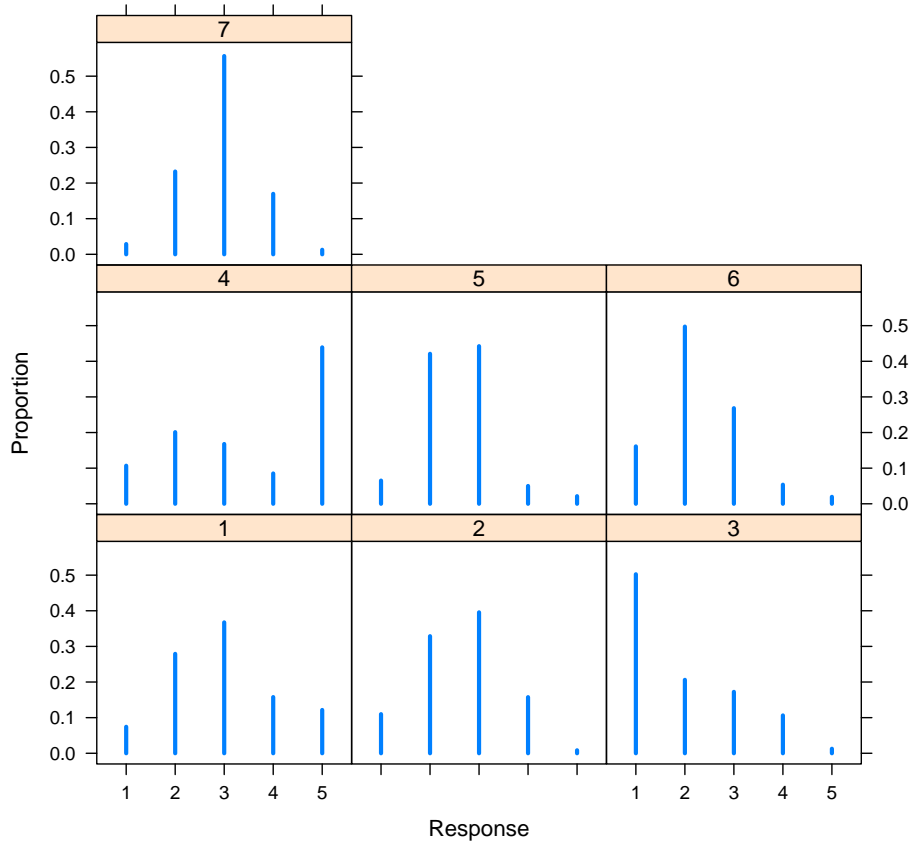
Make_Parents_Proud



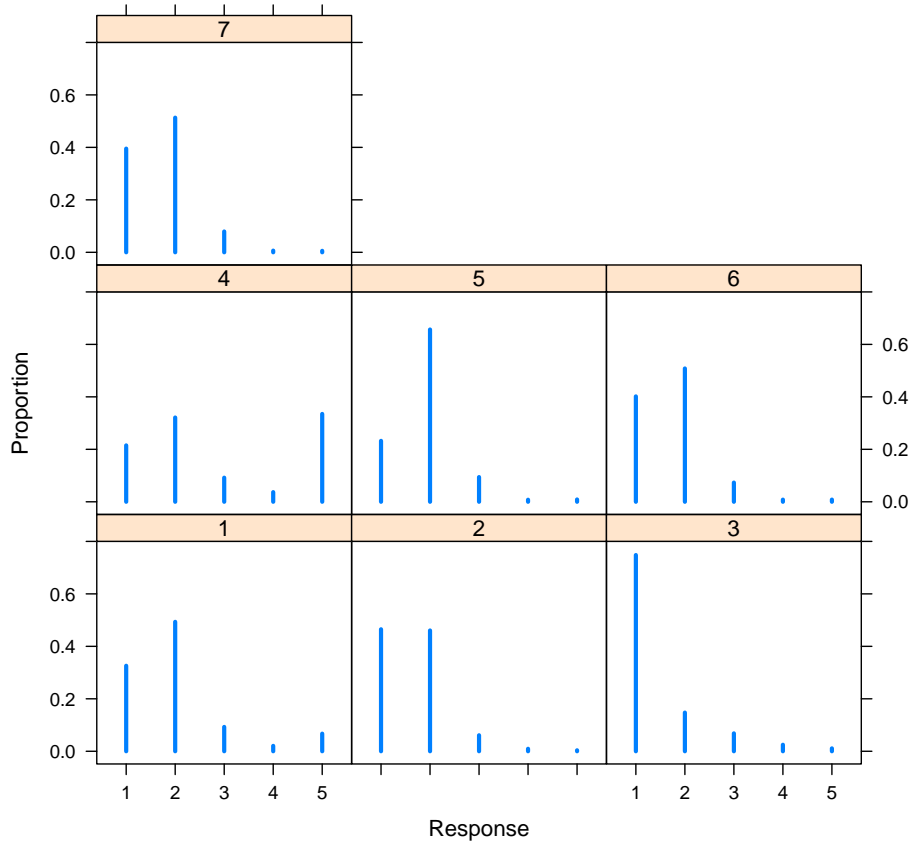
Be_Self



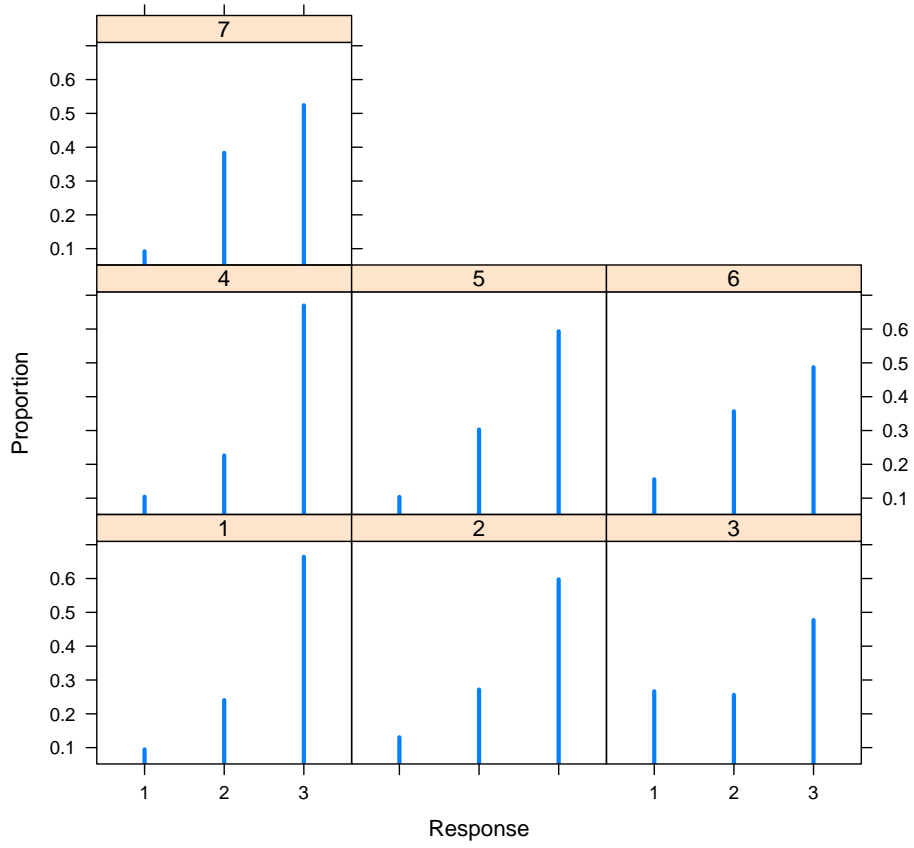
Friends_Expect



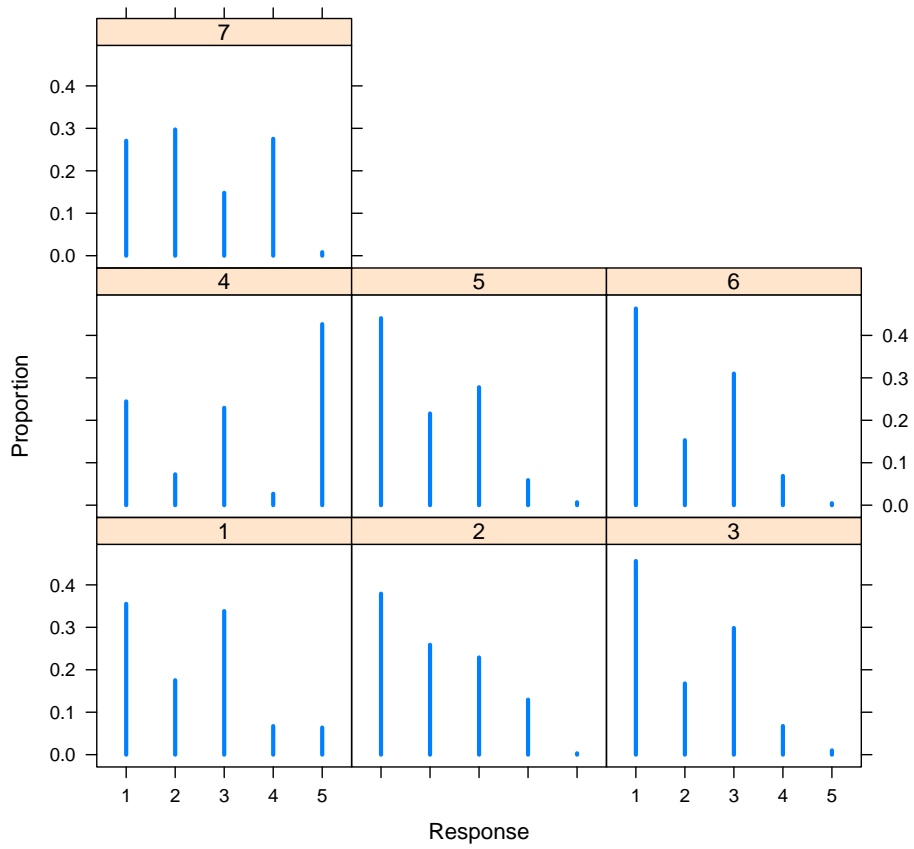
Self_Selected_Goals



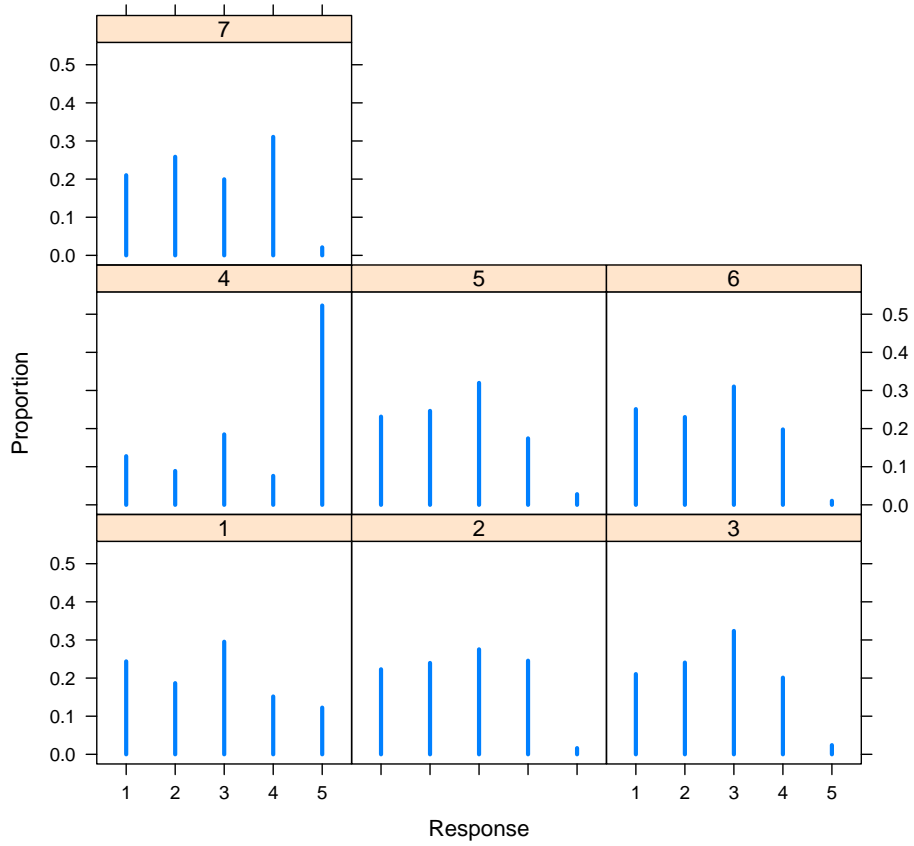
Politics_Important



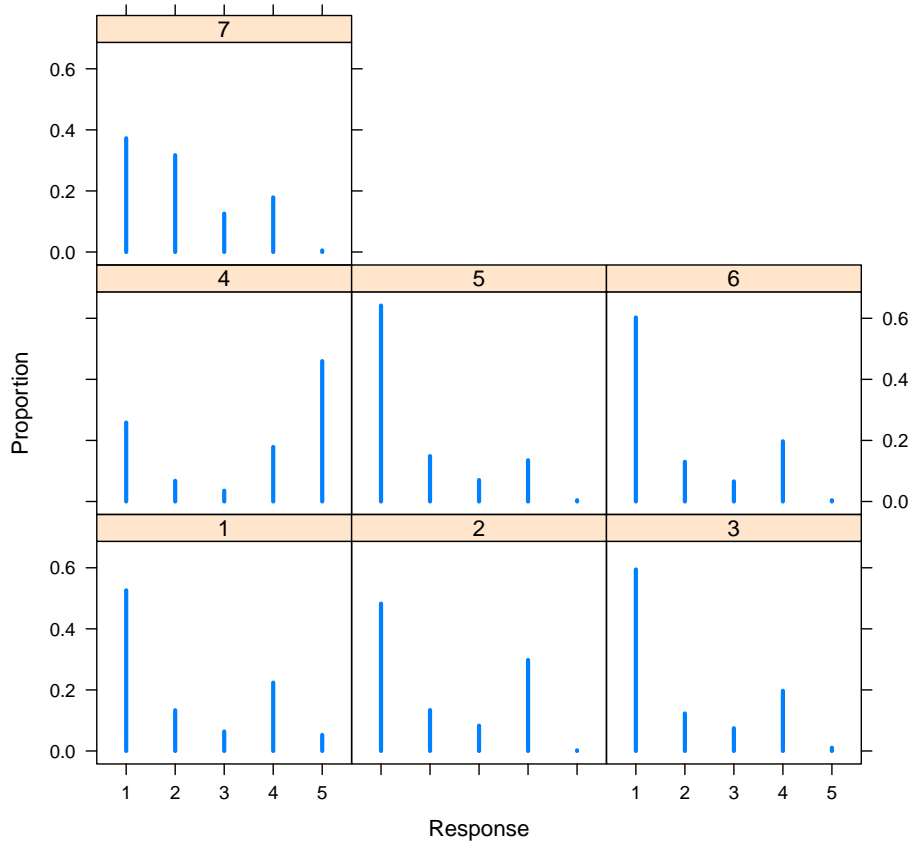
National_Priority_1



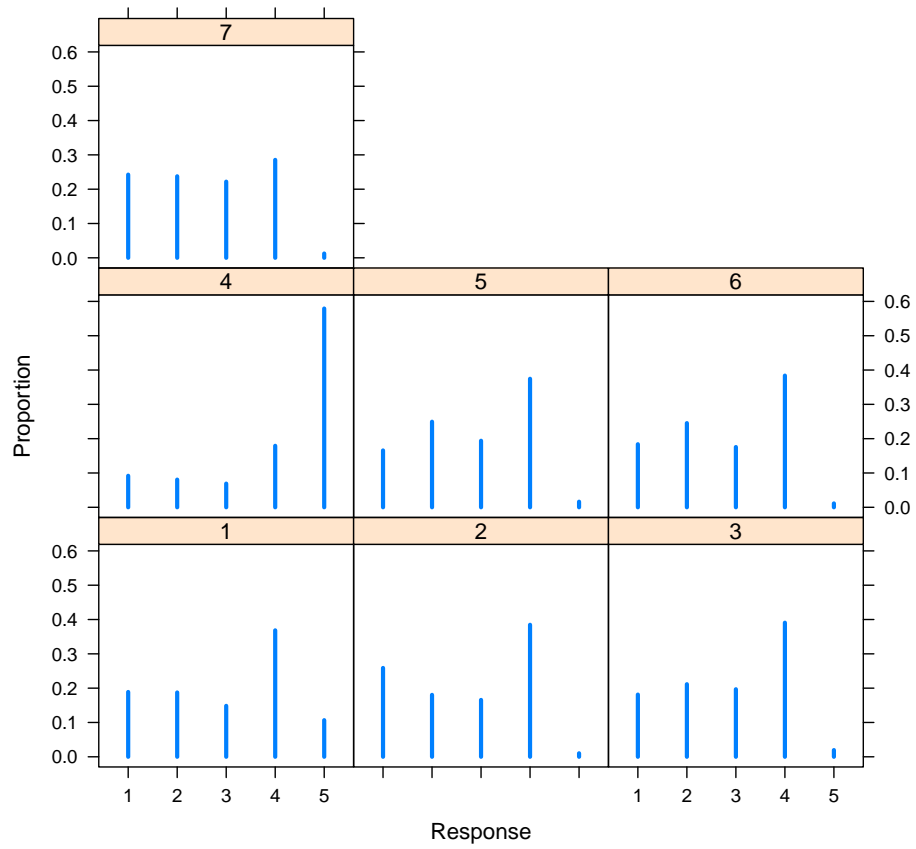
National_Priority_2



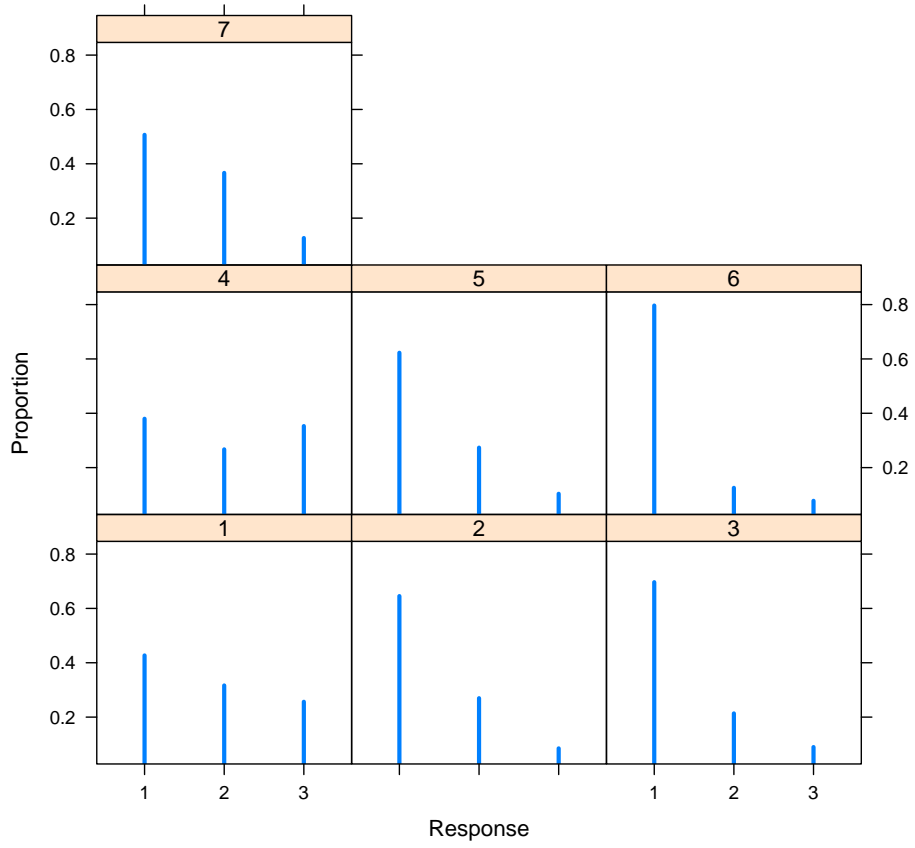
Most_Important_Problem_1



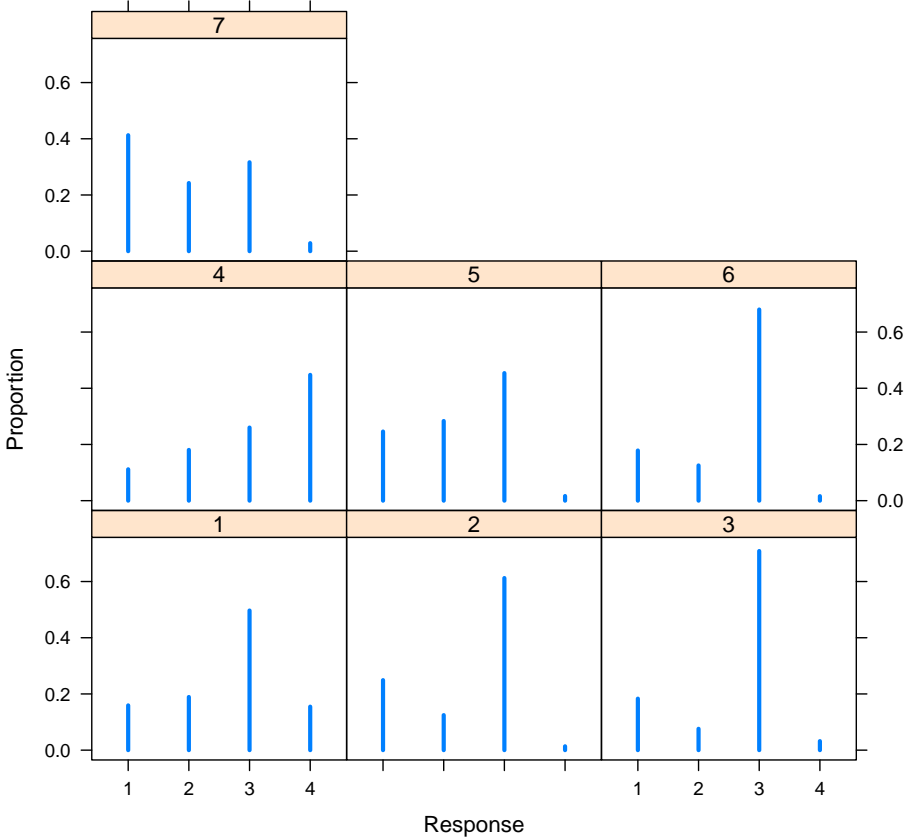
Most_Important_Problem_2



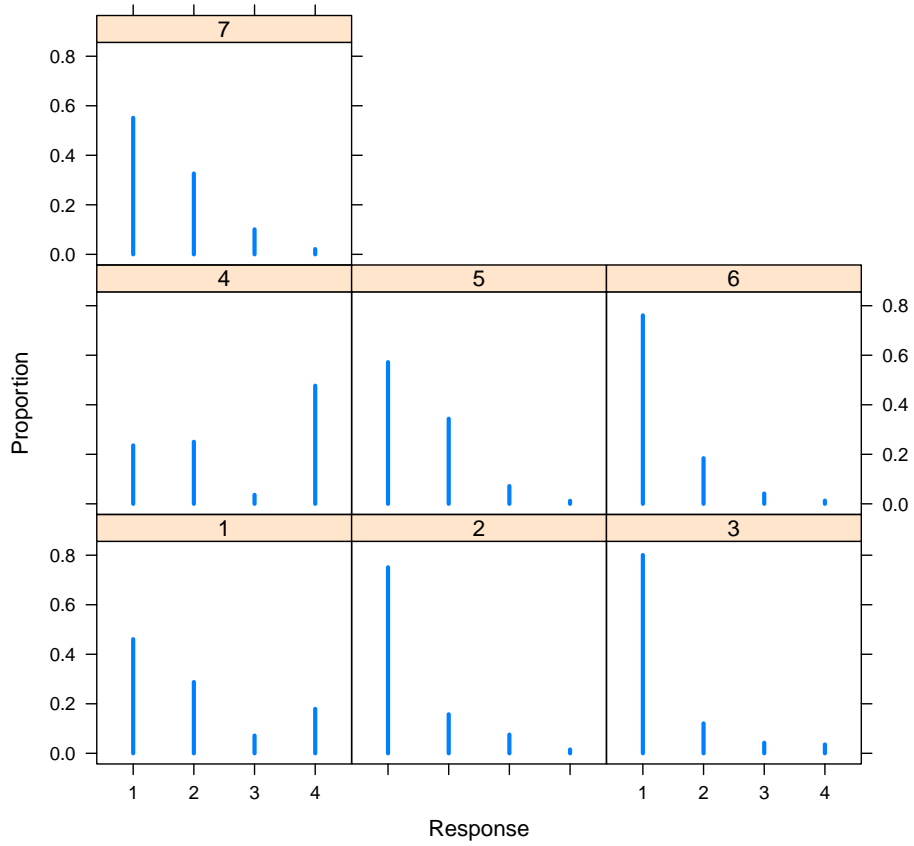
Fight_War



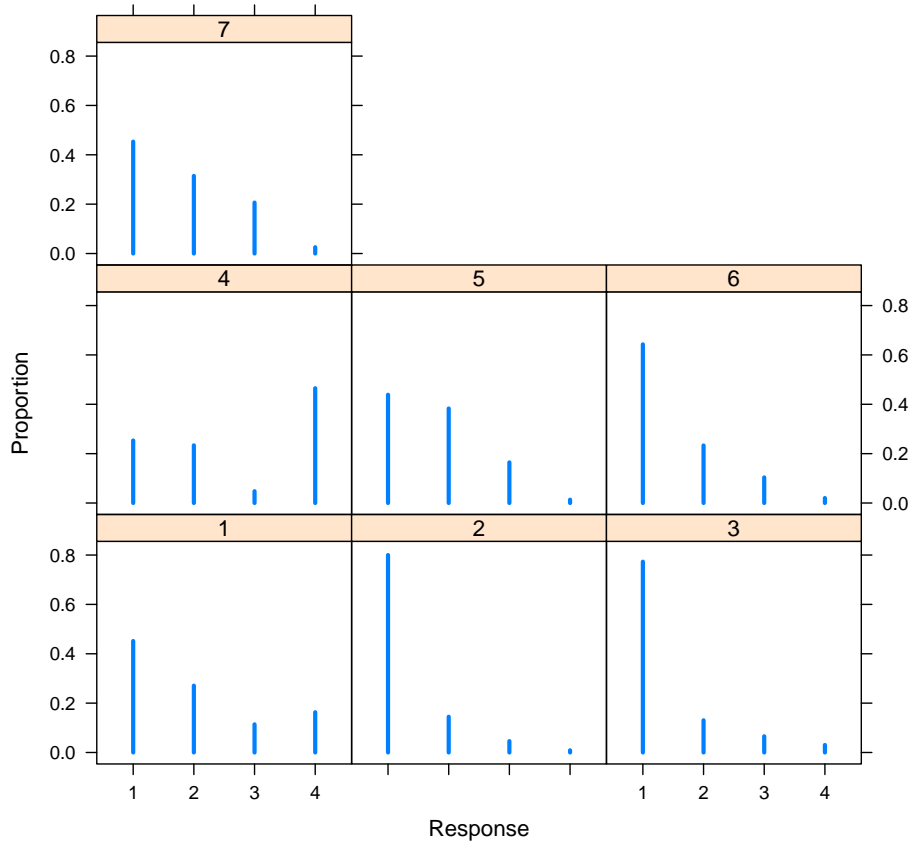
Less_Important_Work



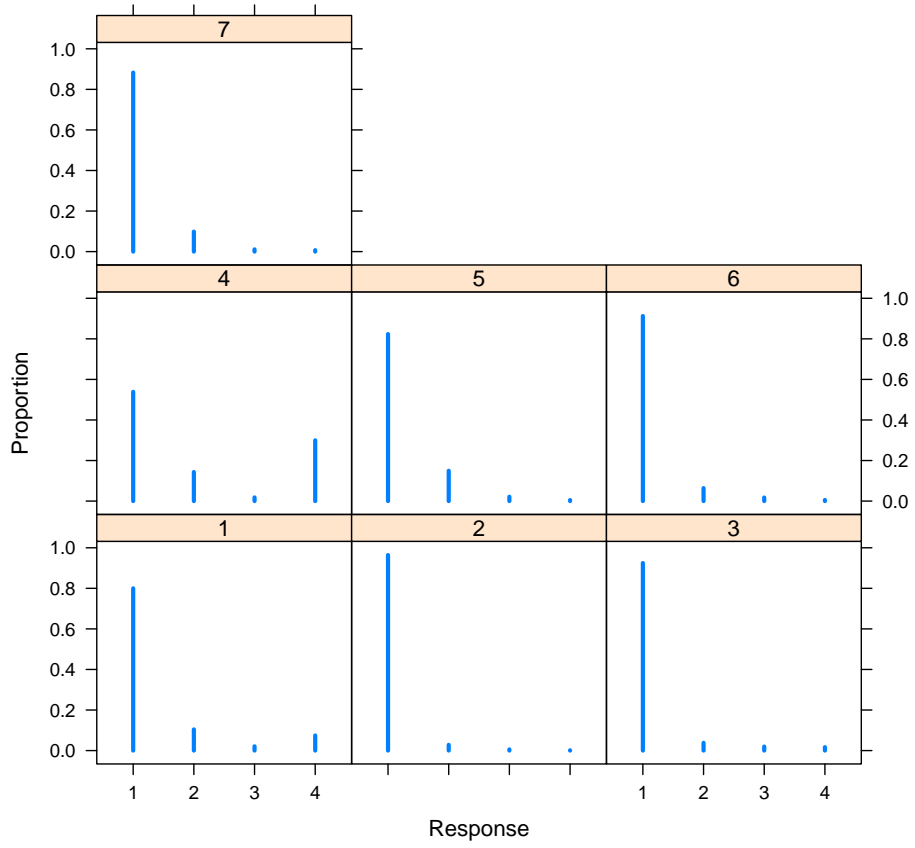
More_Emphasis_Technology



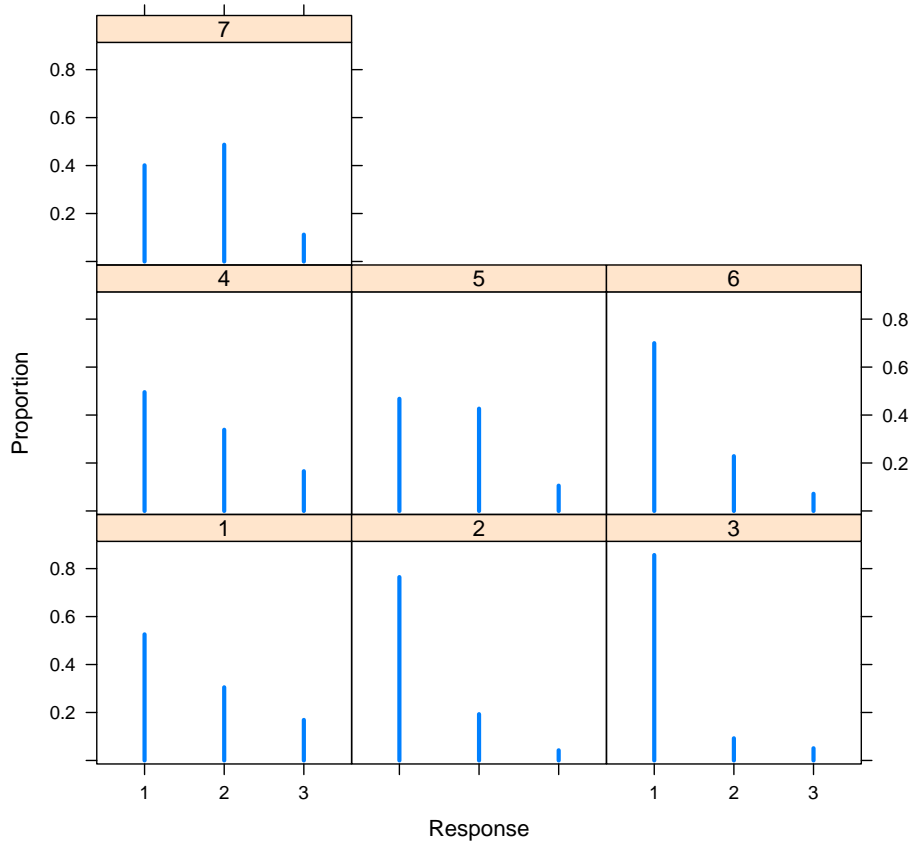
More_Respect_Authority



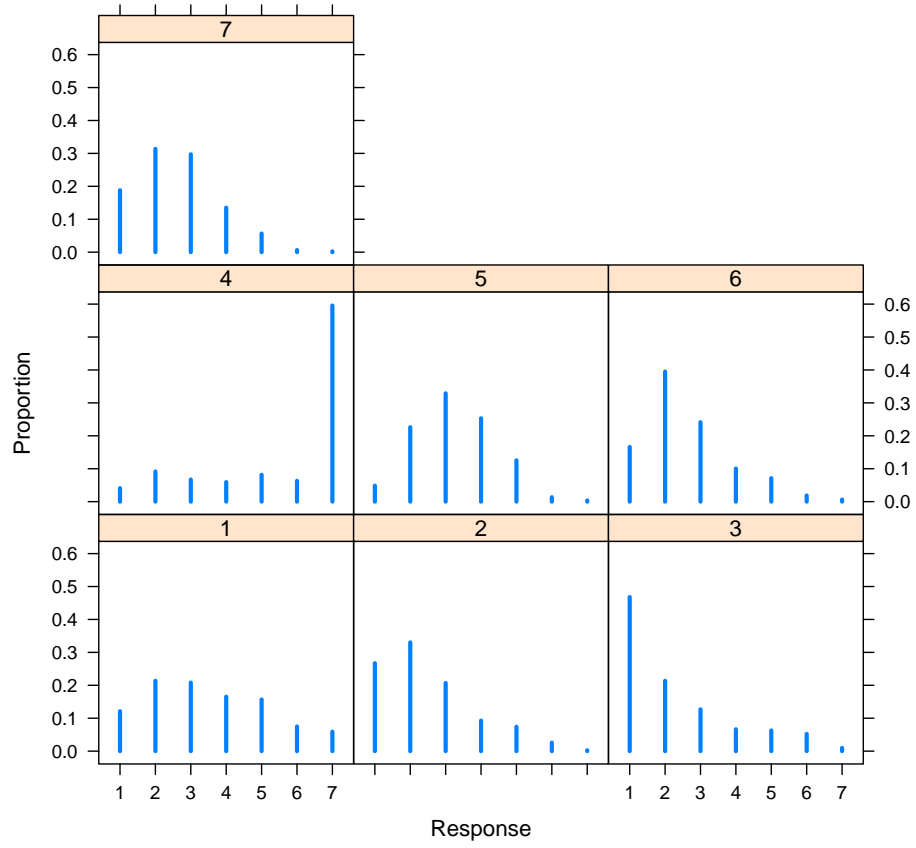
More_Emphasis_Family



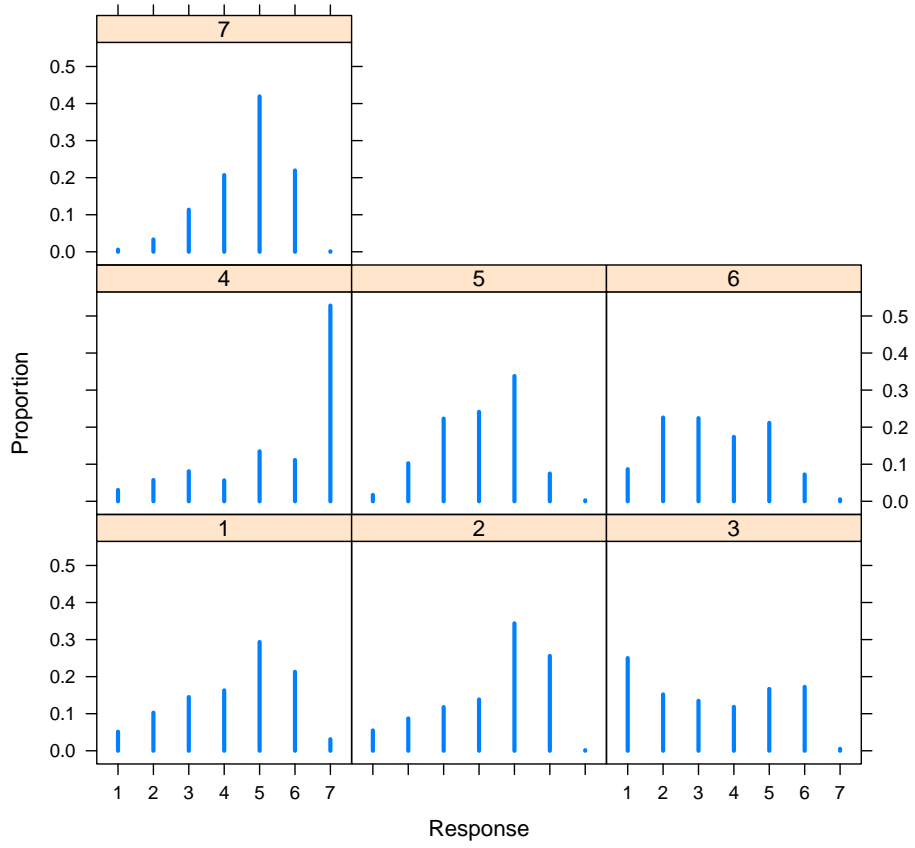
Work_Important



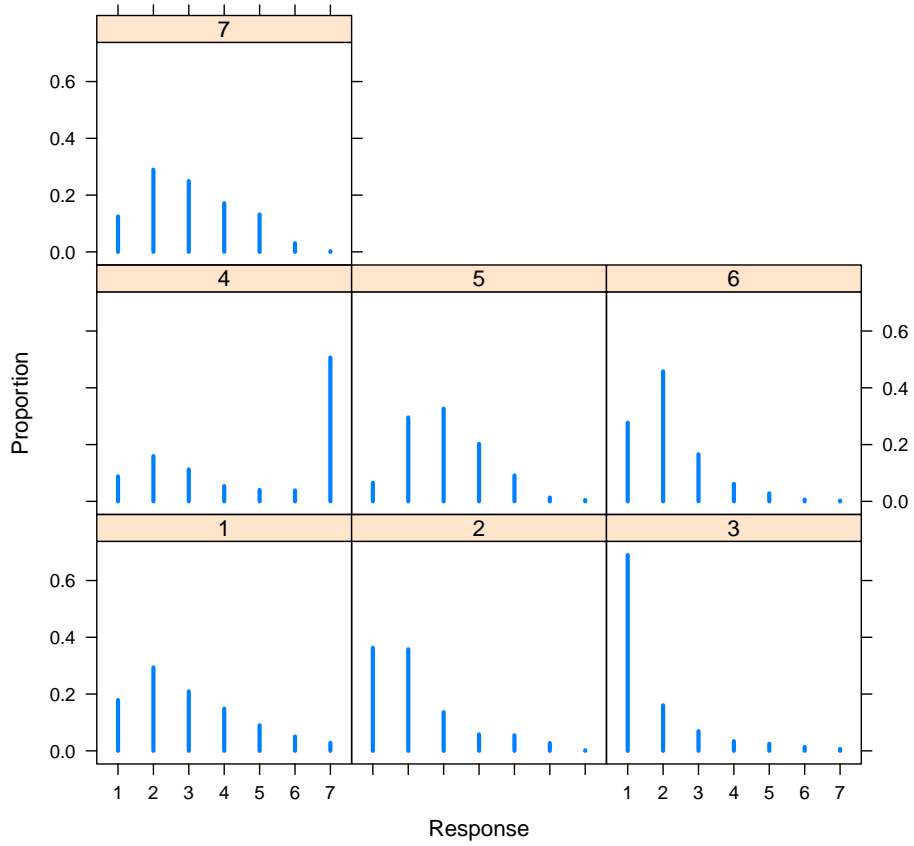
Think_New_Ideas



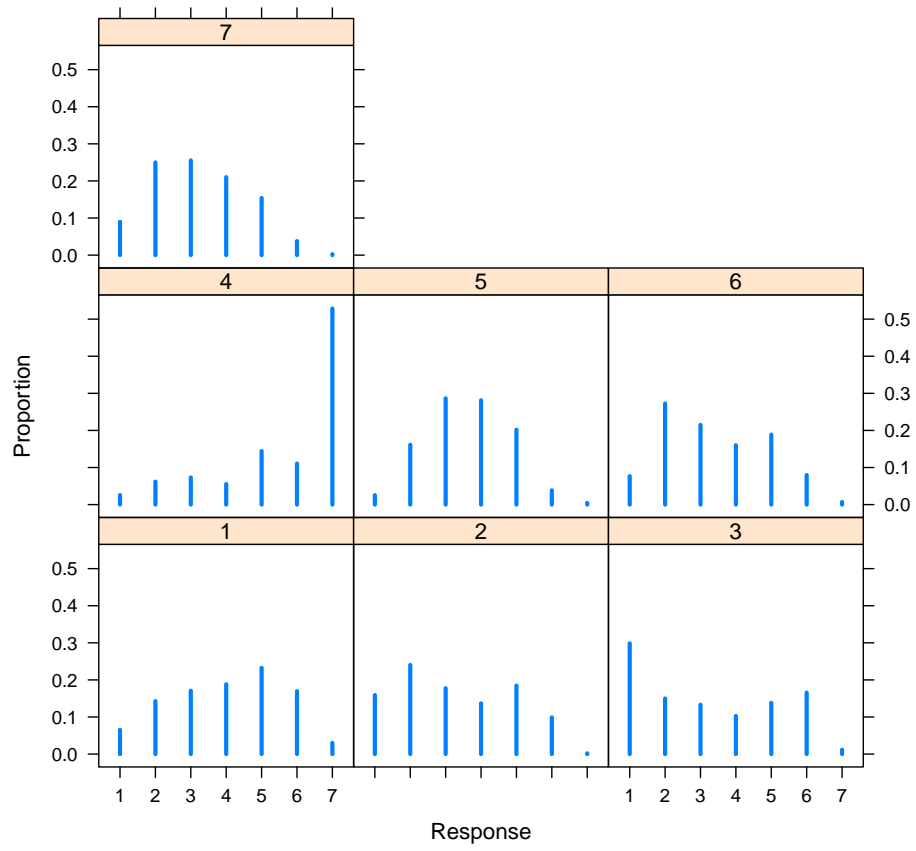
Be_Rich



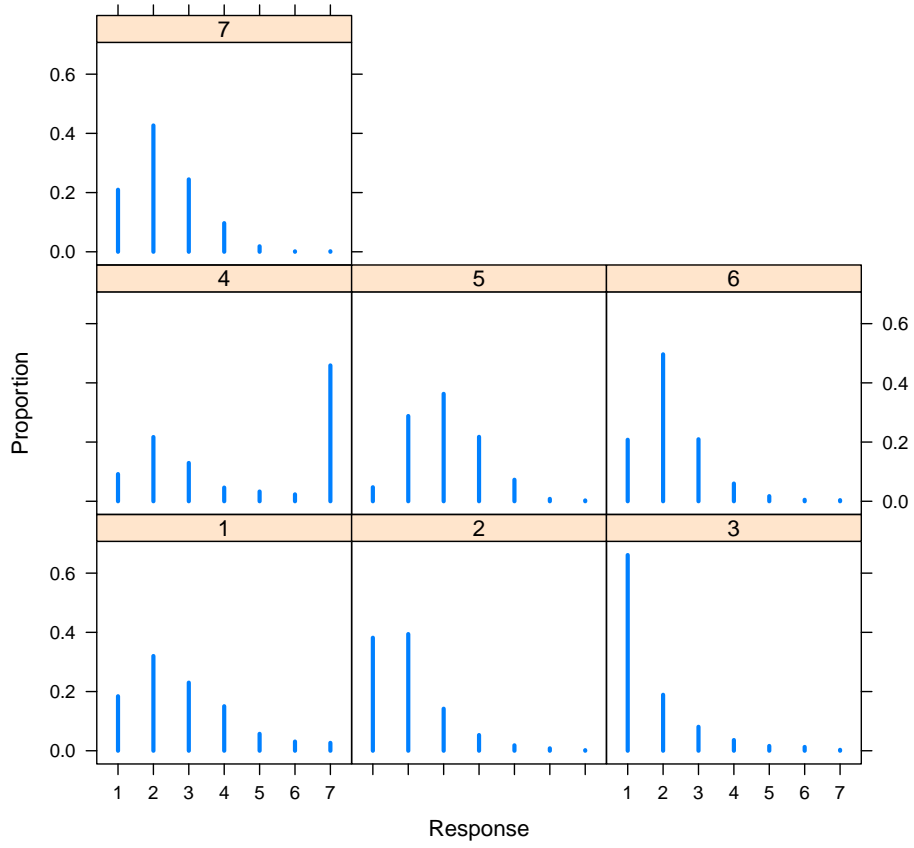
Secure_Surroundings



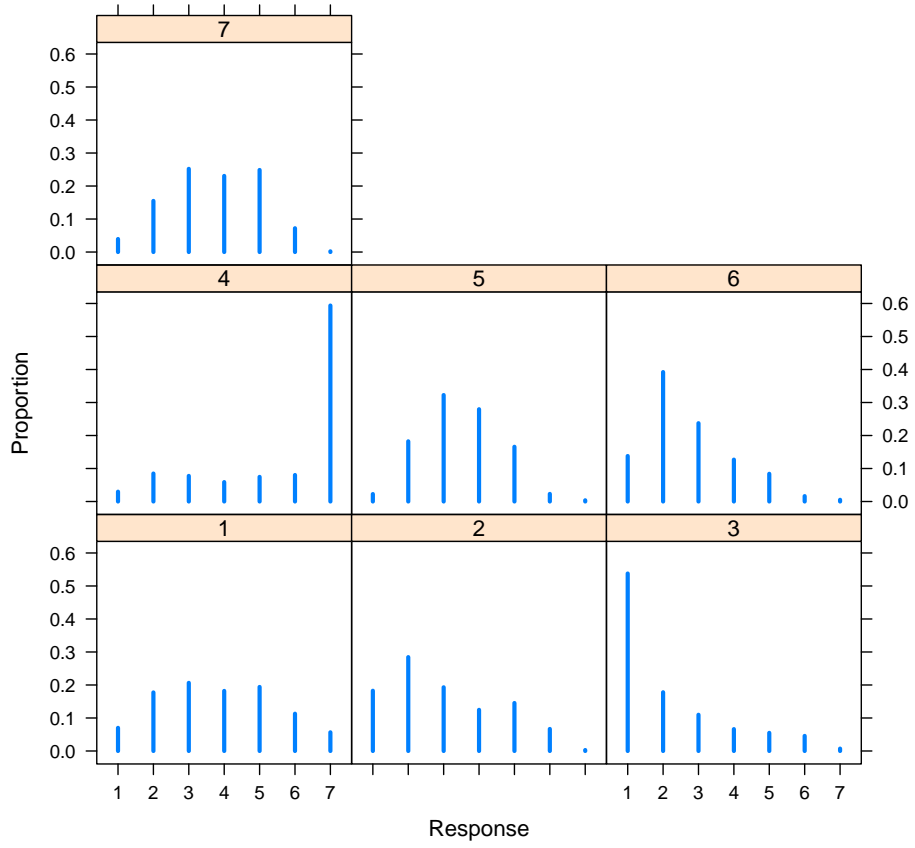
Have_Good_Time



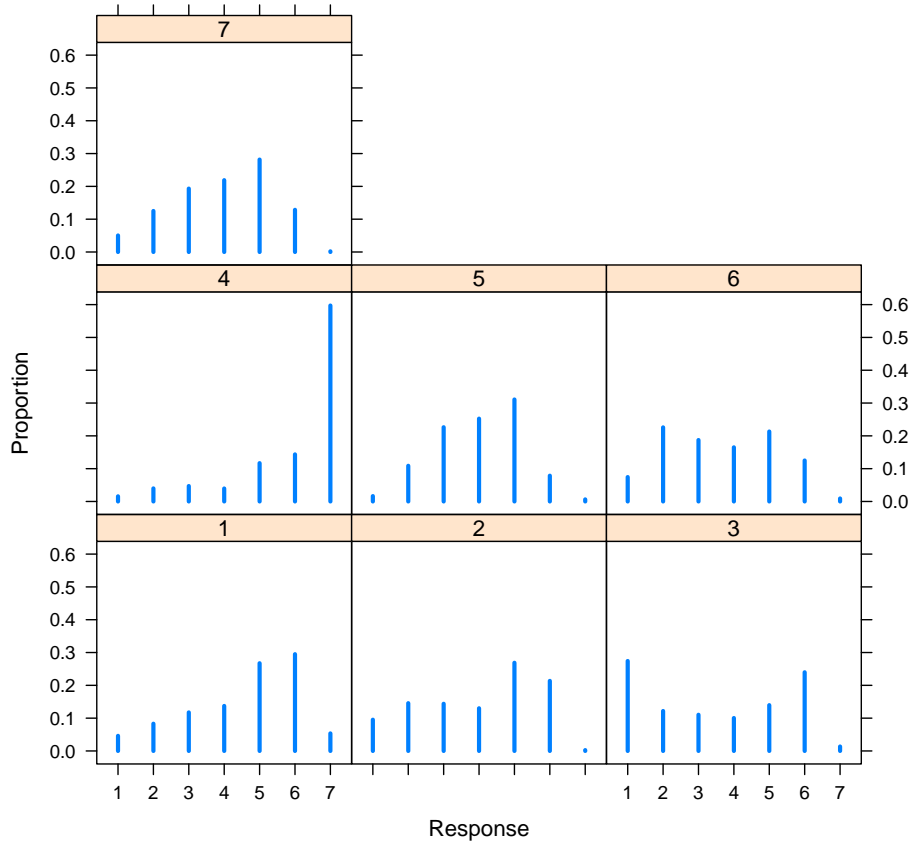
Help_People



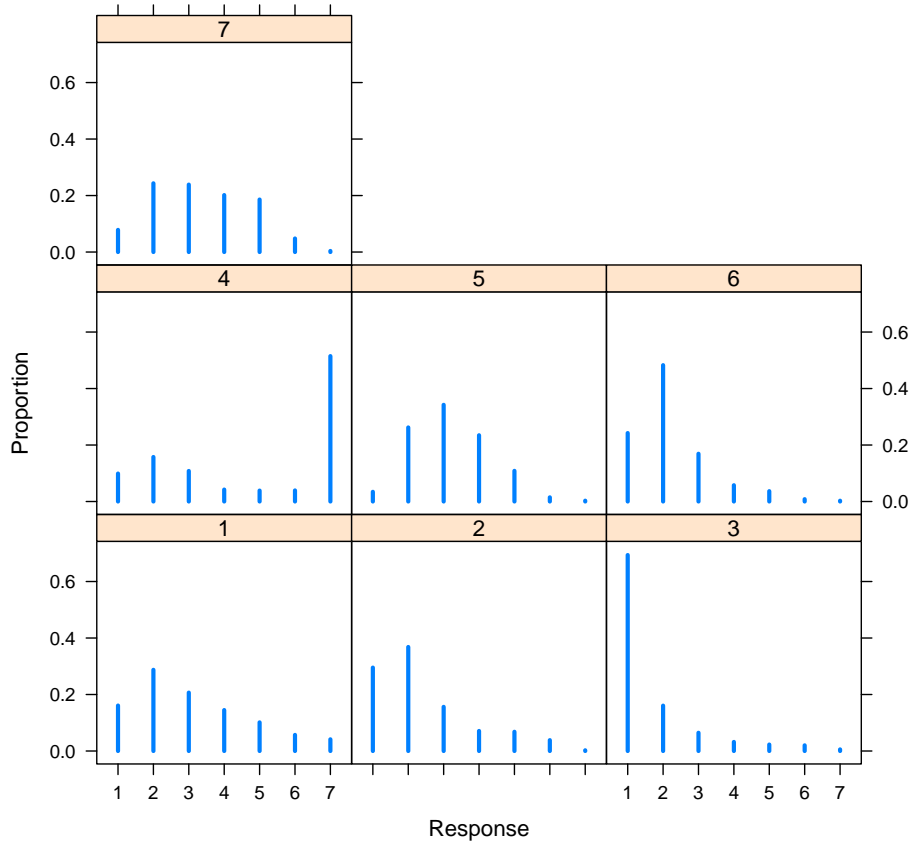
Very_Successful



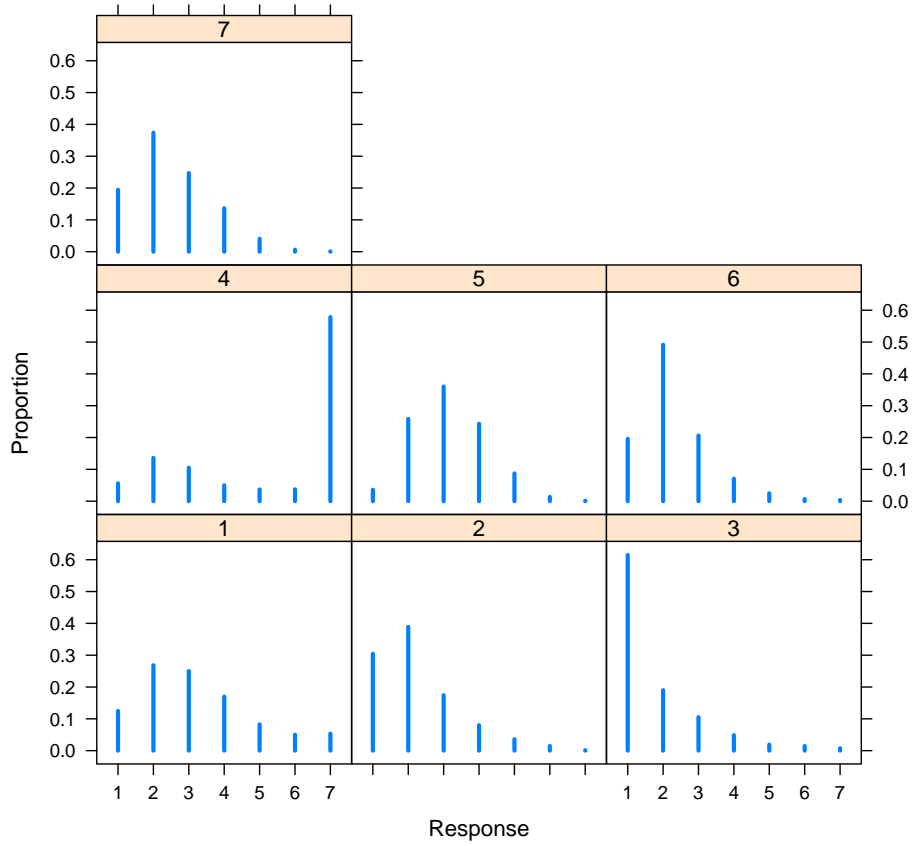
Adventure_Risk_Taking



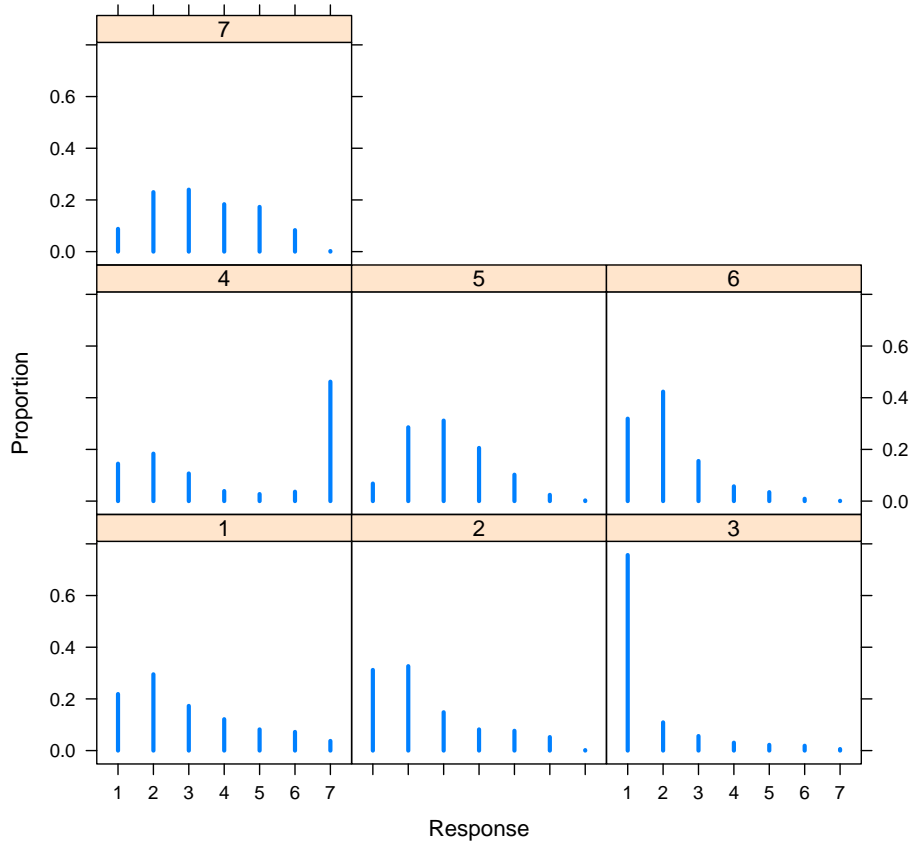
Behave_Properly



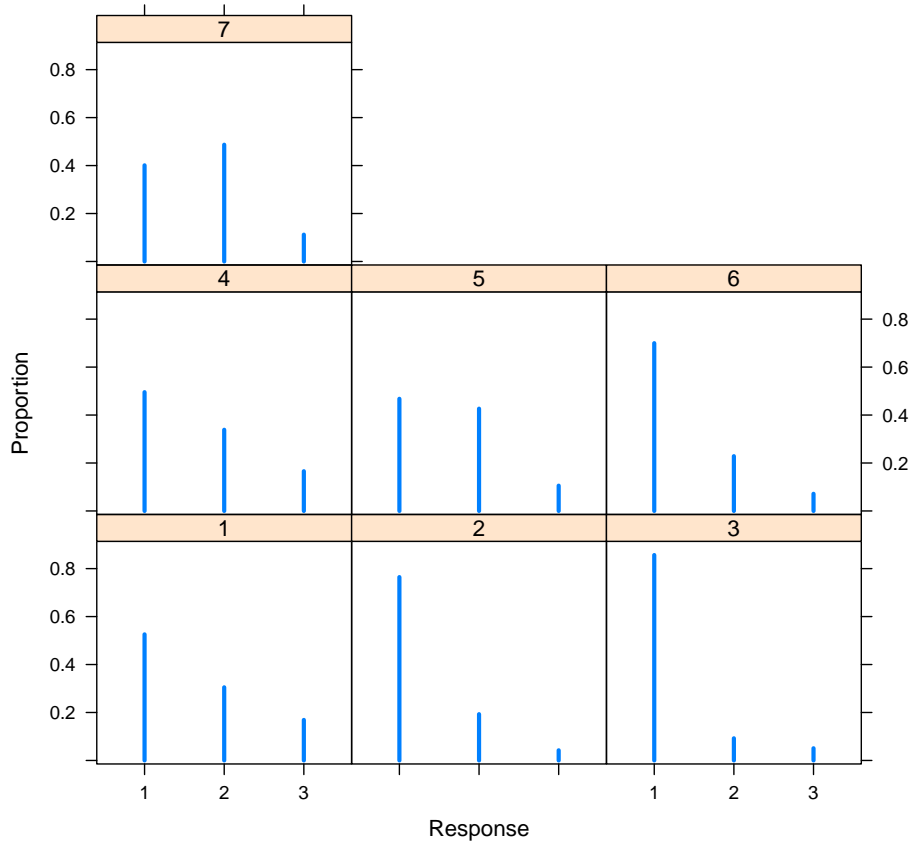
Look_After_Environment



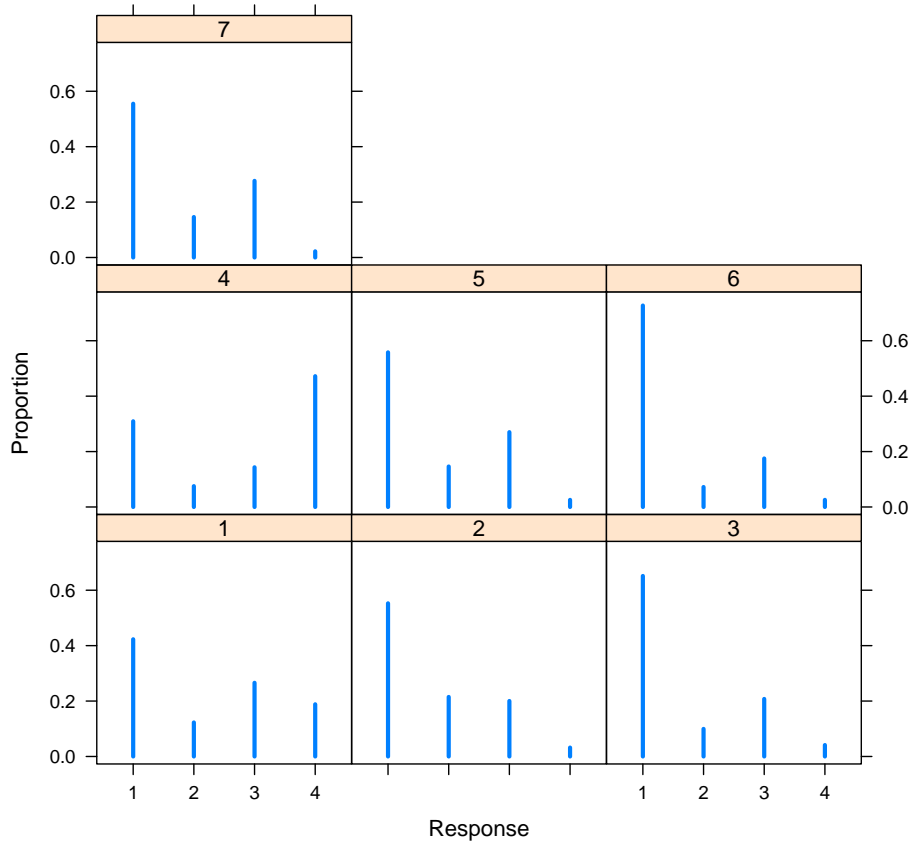
Tradition



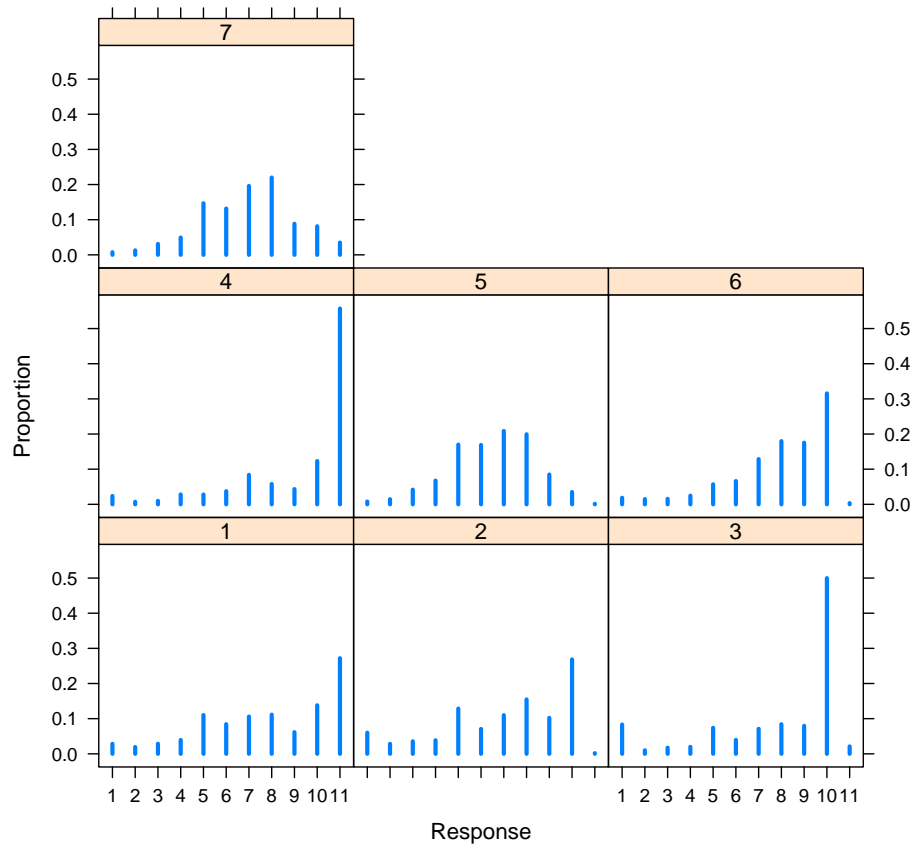
Religion_Important



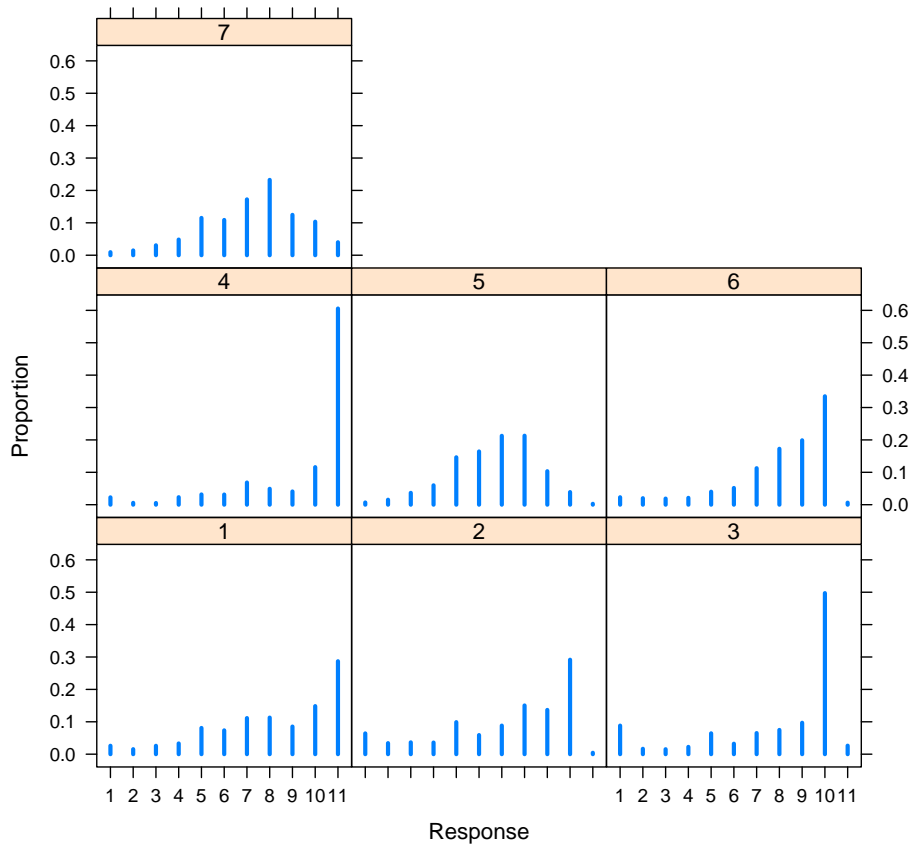
Scientific_Advance



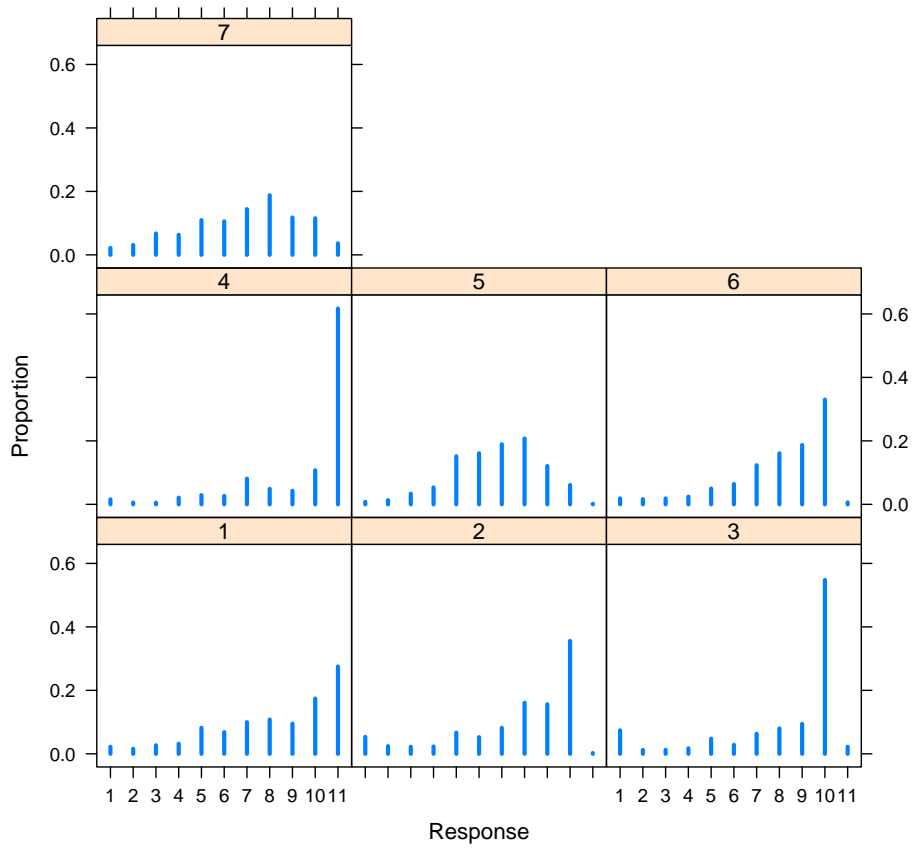
Science_Better



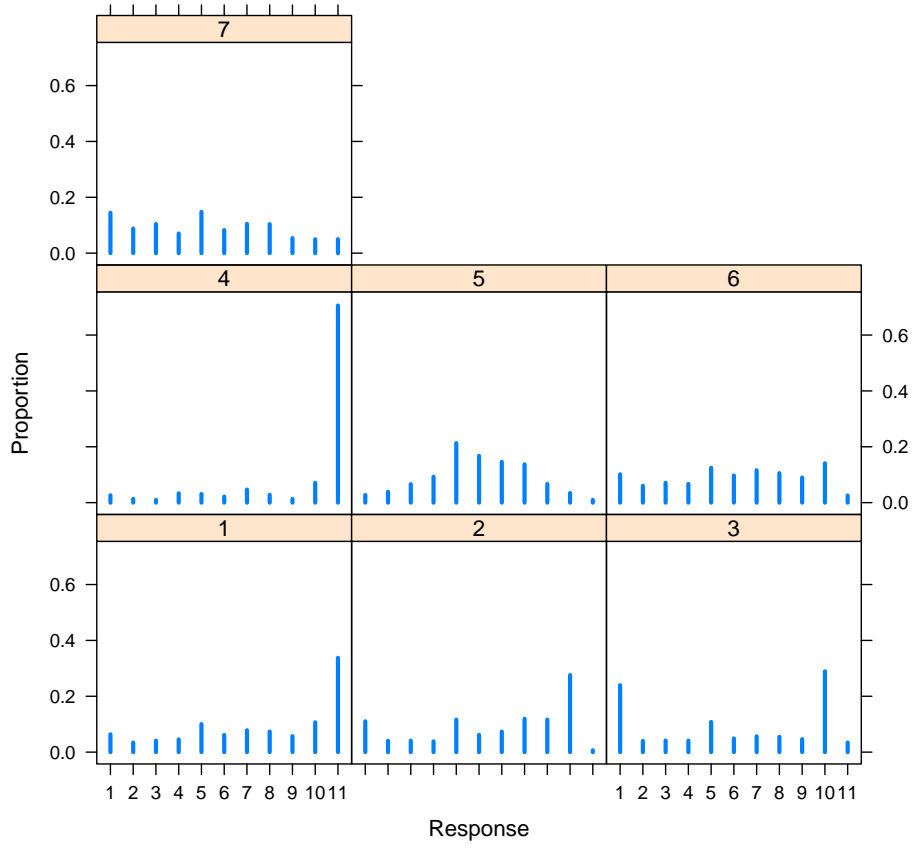
Science_More_Opportunities



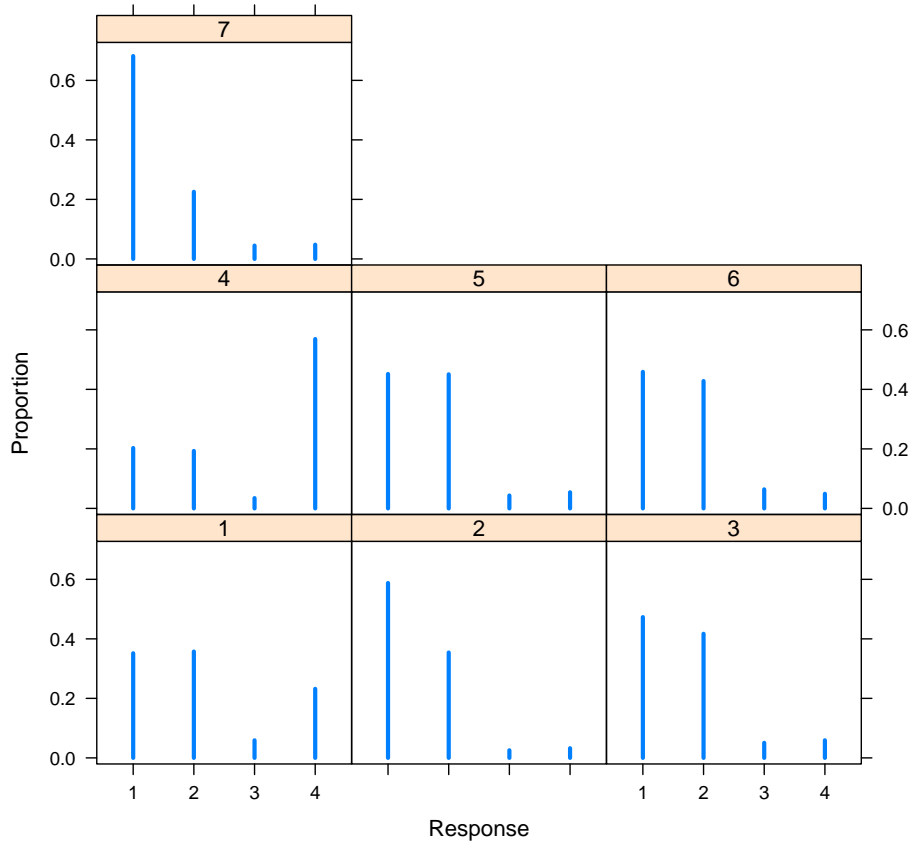
Science_Fast_Changing



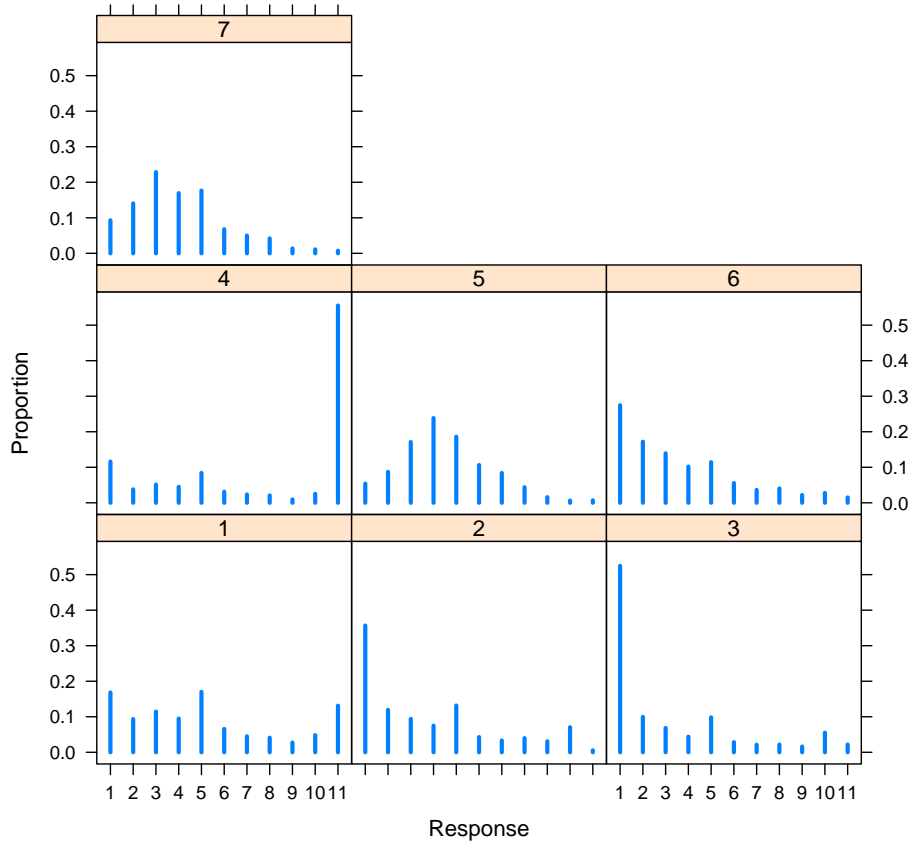
Science_And_Faith



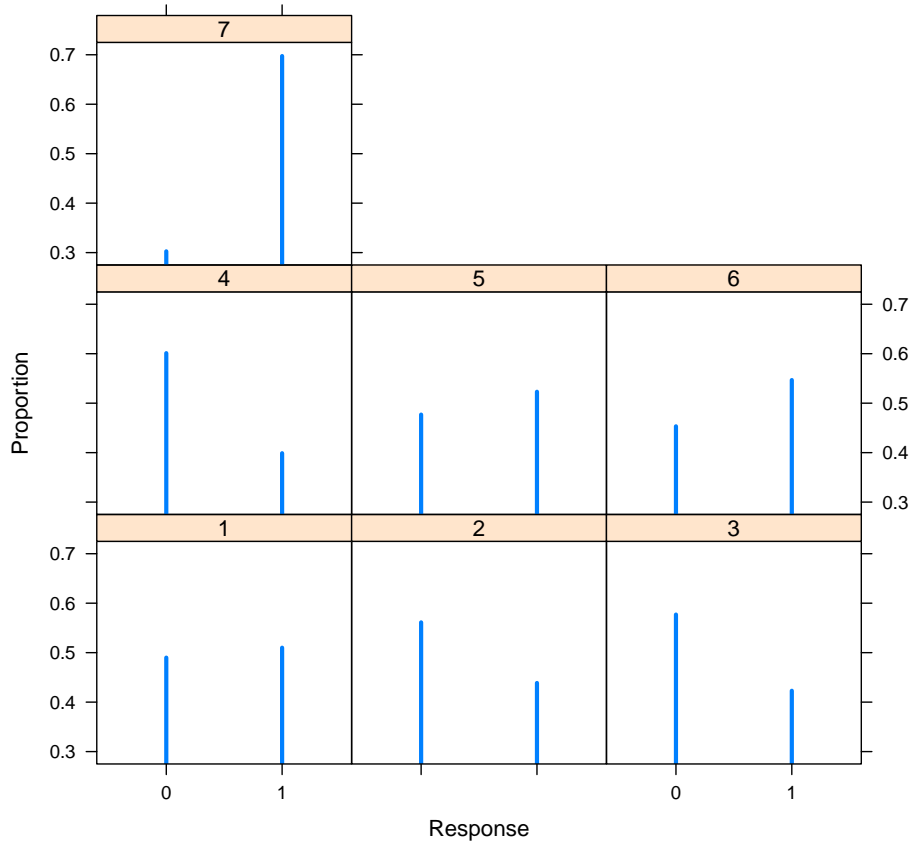
Environmental_Economic_Growth



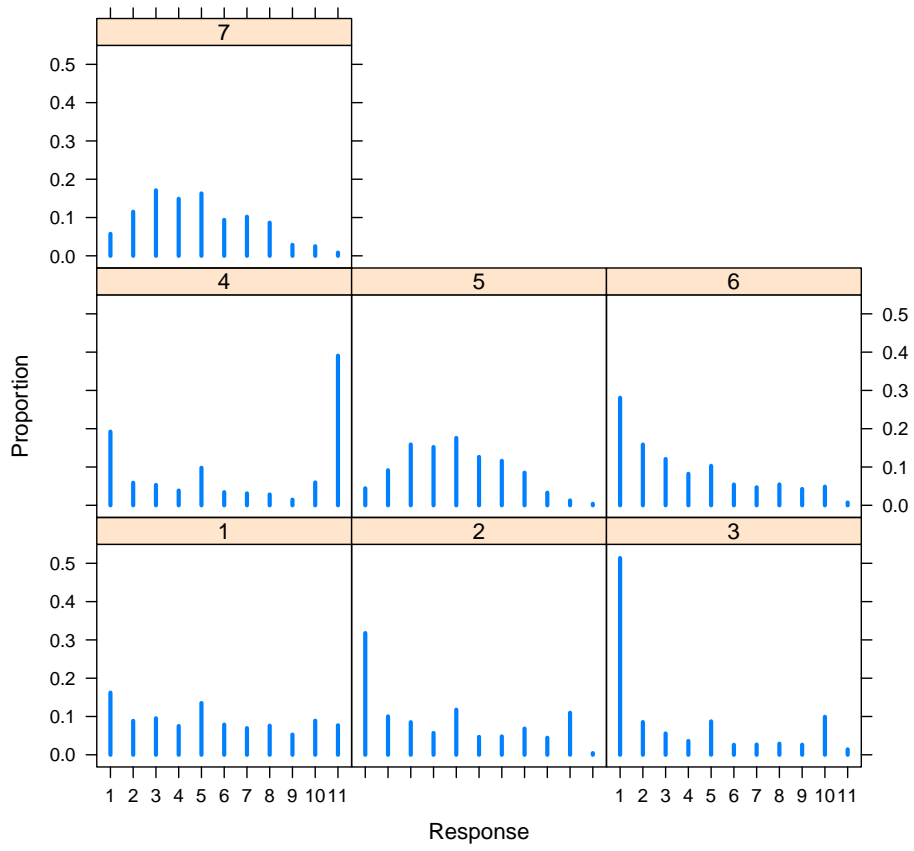
Competition_Good



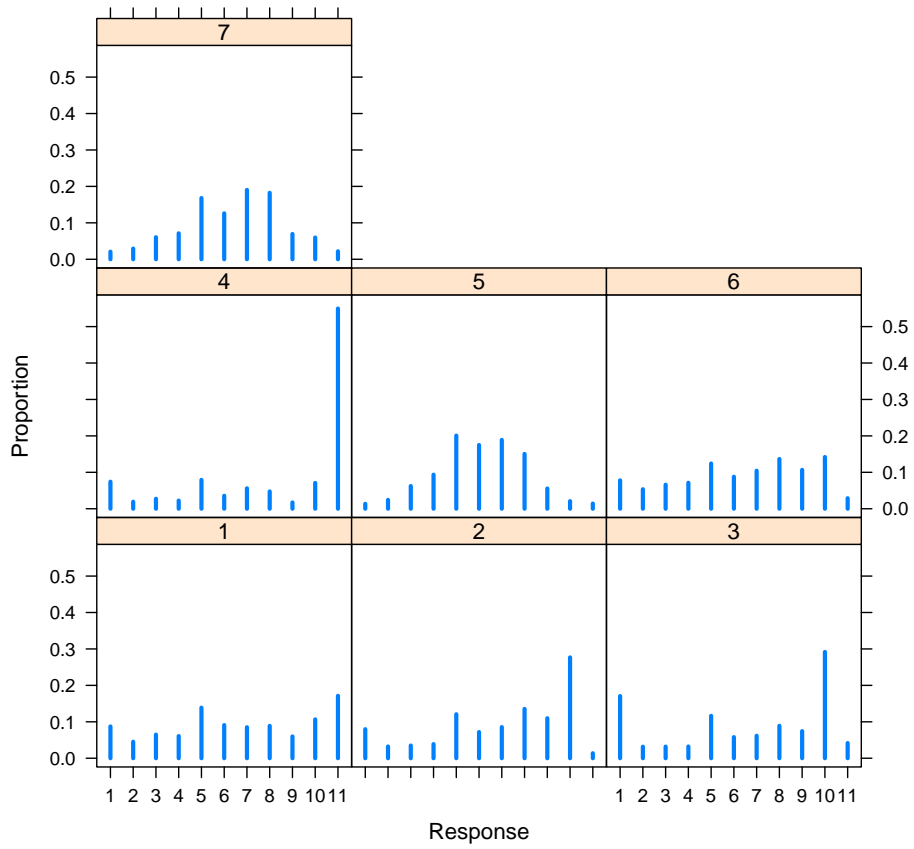
Child_Independence



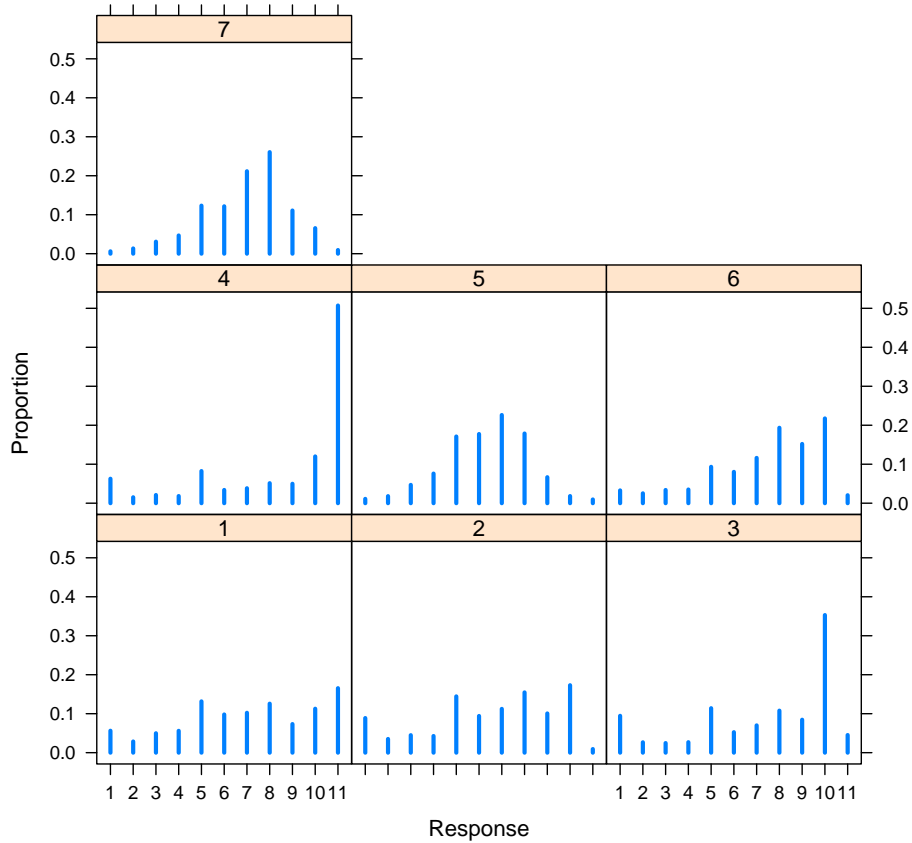
Hard_Work



Wealth_Accumulation



Science_Better_Worse



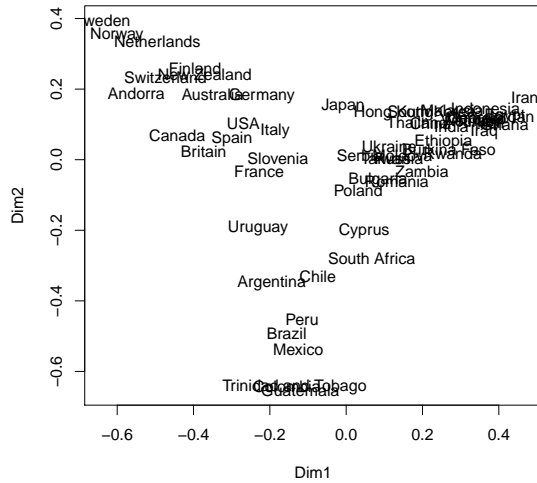
6 Country Specific Face Validity

The previous section showed that the cultures the algorithm identifies are coherent. In this section we characterize country specific variation for a subset of countries. This demonstrates that the categories that we identify align with intuitive understandings of the mix of cultures within a particular country.

Table 1: Country Specific Distribution

Country	1	2	3	4	5	6	7
Sweden	0.01	0.07	0.00	0.00	0.01	0.01	0.89
US	0.01	0.19	0.01	0.02	0.31	0.04	0.41
Iran	0.02	0.00	0.18	0.00	0.03	0.77	0.00
Mexico	0.05	0.71	0.07	0.04	0.05	0.05	0.04

Figure 3: First Two Principal Components of Country Specific SubCulture Distribution



6.1 Distribution of Sweden

The first row of Table 1 provides the distribution of types for Sweden. By far the largest category is category 7, a group that we identify in the main text as mainly "post-materialist". That is, this is a group of people who care less about the distribution of resources, are less responsive to authority, and are more worried about their friends. This aligns, broadly, with identified trends in Scandinavian countries.

The only other group of note is the second category. This group is more conservative but still tolerant. As the above figures show, this is a group open to having a variety of neighbors, but still believe religion is important. As we show below, this is broadly consistent with immigrant attitudes in the country.

6.2 Distribution of US

The United States is a mix of three primary categories. Like Sweden, the US has a substantial "post-materialist" cluster, but an even larger share of its population is in category 2.

The United States also has a large share of Category 5. This category describes respondents who tend to be focused on material issues. They care more about a stable economy than making a city beautiful. This is a group focused on the material concerns primarily, while subordinating other choices. This describes a large share of Americans who identify as pragmatic and traditional in their policy priorities.

6.3 Distribution of Iran

Iran differs sharply from the US and Sweden. In Iran members of category 3 and 6 comprise the largest share of the population. Those two categories are extremely similar, as we now detail here. As we note in the main text, members of category 6 are willing to fight wars and object to having

neighbors who have AIDS, are homosexual, or are unmarried. This reflects a set of respondents who are closed minded and fearful from those who are perceived as different.

Category 3 respondents care a great deal about safety, are highly religious (even for children) and are discriminatory towards out-groups. This would reflect the most religiously conservative elements within Iran.

6.4 Distribution of Mexico

Unlike Iran, respondents from Mexico are much more tolerant. By and large the respondents from Mexico come from category 2: a category that, as we mentioned above, is much more pragmatic and focused on improving material circumstances, while still believing religion is important.

7 Calculating Distance Between Groups

To calculate the distance between types we compare the question-specific expected distributions for each type. Specifically, for type k call the expected distribution over responses for question r $\hat{\theta}_{kr}$, which we can calculate directly from the estimated model. Using this distribution across questions, we can calculate the distance between groups. Specifically, for group k and k' we will define the distance, $d_{k,k'}$,

$$d_{k,k'} = \sum_{r=1}^R |\hat{\theta}_{kr} - \hat{\theta}_{k'r}|$$

Calculating the distance this way ensures that our comparisons capture how members of each group answer questions differently.

8 Relation to World Cultures

To what extent do these cultural categories match up to scholarly conceptualizations of differences between world cultures? Huntington (1996) argues that the world can be divided into a series of world civilizations (i.e., Western, Orthodox, African, Latin American, Islamic, Buddhist, etc). From a cultural values perspective, Huntington’s typology only holds up in a very limited way and not in the manner he predicted. Latin American countries tend to be quite homogenous with regard to stated beliefs and values. Muslim-majority countries seem to be predominantly split between categories 3 and 6, which are relatively similar to each other but differ on the dimension of women’s rights. Most societies Huntington categorizes as “Western” appear to be a mix of category 7 values and the values of a variety of subcultures. The countries in the former Soviet-influenced space exhibit strikingly little cultural homogeneity. Most East Asian countries have large clusters of citizens with category 5 values; it is not clear that these values parallel Buddhist teachings and it is notable that there is considerable variation in cultural values across China.

One dimension on which Huntington (1996) may enjoy some support relates to the *distance* between what might be called “Western” values and the cultural values of Muslim-majority societies. Although there is considerable diversity within the Islamic world, the world cultures that appear to be most dissimilar are 3 — typically found of Egypt, Iraq and Jordan — and 7 — typically found in Finland, Sweden and, to a lesser degree, the United States.

References

- Bishop, Christopher. 2006. *Pattern Recognition and Machine Learning*. Springer.
- Grimmer, Justin. 2011. “An Introduction to Bayesian Inference via Variational Approximations.” *Political Analysis* 19(1):32–47.
- Hastie, Trevor, Robert Tibshirani and Jerome Friedman. 2001. *The Elements of Statistical Learning*. Springer.
- Huntington, Samuel. 1996. *The Clash of Civilizations and the Remaking of World Order*. New York, NY: Simon & Schuster.
- Minka, Thomas. 2000. “Estimating the Parameters of a Dirichlet.” Mimeo.