Missing Data Issues in Large-Scale Surveys: Questionnaire Rotation and Data Fusion

David Kaplan

Department of Educational Psychology

Invited Talk to the Joint Research Center of the European Commission
This talk considers two important issues relevant to the design and analysis of international large scale assessments.

1. Rotation and imputation of context questionnaires.
   - A work in progress.

2. Data fusion (also referred to as statistical matching)

Both issues have the theme of missing-data-by-design in common.
My experience with large-scale assessments

- Member: PISA Questionnaire Expert Group 2004–Present
- Chair: Questionnaire Expert Group for PISA 2015, 2013–Present
- Member: NAEP Design and Analysis Committee
- Member: NAEP Questionnaire Standing Committee
Interest in the rotation of context questionnaires (CQ) is motivated by the increase in the amount of contextual knowledge that can be used to support policy analysis.

PISA 2012 implemented a rotation of the CQ that contained 25 common items and which increased the amount of additional information covered by about 1/3.
### Table 6.4

Student questionnaire: Rotated forms A, B and C

<table>
<thead>
<tr>
<th>Form A</th>
<th>Form B</th>
<th>Form C</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q. n°</td>
<td>Description</td>
<td>Q. n°</td>
</tr>
<tr>
<td>ST01-28</td>
<td>Common Part (see Table 6.3)</td>
<td>ST01-28</td>
</tr>
<tr>
<td>ST129</td>
<td>Intrinsic and Instrumental Motivation for Mathematics</td>
<td>ST42</td>
</tr>
<tr>
<td>ST135</td>
<td>Subjective Norms</td>
<td>ST77</td>
</tr>
<tr>
<td>ST137</td>
<td>Mathematics Self-Efficacy</td>
<td>ST79</td>
</tr>
<tr>
<td>ST143</td>
<td>Perceived Control of Mathematics Performance</td>
<td>ST80</td>
</tr>
<tr>
<td>ST144</td>
<td>Attribution to Failure in Mathematics</td>
<td>ST81</td>
</tr>
<tr>
<td>ST146</td>
<td>Mathematics Work Ethics</td>
<td>ST84</td>
</tr>
<tr>
<td>ST148</td>
<td>Mathematics Interests</td>
<td>ST83</td>
</tr>
<tr>
<td>ST149</td>
<td>Mathematics Behaviour</td>
<td>ST84</td>
</tr>
<tr>
<td>ST153</td>
<td>Persistence</td>
<td>ST85</td>
</tr>
<tr>
<td>ST154</td>
<td>Openness for Problem Solving</td>
<td>ST86</td>
</tr>
<tr>
<td>ST156</td>
<td>Problem-Solving Strategies (ST)</td>
<td>ST87</td>
</tr>
<tr>
<td>ST158</td>
<td>Problem-Solving Strategies (ST)</td>
<td>ST88</td>
</tr>
<tr>
<td>ST158</td>
<td>Problem-Solving Strategies (ST)</td>
<td>ST89</td>
</tr>
<tr>
<td>ST163</td>
<td>Learning Strategies (Control vs. Elaboration vs. Memorisation)</td>
<td>ST91</td>
</tr>
<tr>
<td>ST164</td>
<td>Perceived Control of Success in School</td>
<td>ST92</td>
</tr>
<tr>
<td>ST167</td>
<td>Mathematics Self-Efficacy</td>
<td>ST93</td>
</tr>
<tr>
<td>ST169</td>
<td>Mathematics Work Ethics</td>
<td>ST94</td>
</tr>
<tr>
<td>ST170</td>
<td>IP of Class Period: Per Week</td>
<td>ST96</td>
</tr>
<tr>
<td>ST171</td>
<td>IP of All Class Periods: Per Week</td>
<td>ST97</td>
</tr>
<tr>
<td>ST172</td>
<td>Mathematics Teacher’s Classroom Management</td>
<td>ST98</td>
</tr>
<tr>
<td>ST183</td>
<td>Perseverance</td>
<td>ST99</td>
</tr>
<tr>
<td>ST184</td>
<td>Mathematics Teacher’s Classroom Management</td>
<td>ST101</td>
</tr>
<tr>
<td>ST185</td>
<td>Mathematics Teacher’s Classroom Management</td>
<td>ST102</td>
</tr>
<tr>
<td>ST186</td>
<td>Mathematics Teacher’s Classroom Management</td>
<td>ST103</td>
</tr>
</tbody>
</table>
In education LSAs, it is typical to rotate the battery of test items to obtain broad content coverage while minimizing student burden.

As with the rotation of the test items in surveys such as PISA and NAEP, the rotation of the context questionnaires presents a massive missing data problem.

Adams, Leitz, & Berezner (2013) have argued that rotation of the CQ is not detrimental to plausible value estimation of latent proficiency based on the full conditioning and population model approach.

Others disagree. Arguably, the jury is still out.
Multiple approaches can be taken to deal with this problem.

1. Don’t use a rotated CQ design at all.

2. Impute CQ and PVs at the same time.

3. Impute CQ first followed by using the imputed CQ to estimate PVs.

A reasonable concern is the time taken to produce these datasets, as well as the format of the public use dataset.
We concentrate on issues related to the imputation of the CQ.

In practice this would then be followed by the use of the conditioning model for estimating the latent proficiency distributions based on the fully imputed data. See von Davier (2013, in Rutkowski, von Davier, and Rutkowski)

For imputation of the rotated CQ a choice must be made regarding the imputation algorithm that would be used to fill in the missing data.

Not all imputation algorithms provide the same results.

Validity criteria must be established.
We will describe two approaches within the fully conditional specification (chained equation) framework of missing data theory.

These approaches follow a general Bayesian framework for imputation based on the fundamental work of Rubin (1987).

1. Predictive mean matching for all variables.

2. Bayesian linear regression for “continuous” variables and proportional odds logistic regression for categorical variables.
The chained equations approach uses a univariate regression model consistent with the scale of the variable with missing data to provide predicted values of the missing data given the observed data.

Once a variable of interest is “filled-in”, that variable, along with the variables for which there is complete data, is used in a sequence to fill in another variable.

Once the sequence is completed for all variables with missing data, the posterior distribution of the regression parameters are obtained via Gibbs sampling and the process is started again.

The algorithm can run these sequences simultaneously $m$ number of times obtaining $m$ imputed data sets.

This is the method used in the R program “mice“ (van Buuren & Groothuis-Oudshoorn, 2011).
Predictive Mean Matching: mice.impute.pmm()

Let $X_{obs}$ be the predictors with observed data and let $X_{miss}$ be the predictors with missing data on the target variable $y$.

1. Obtain $\hat{\beta}$ based on $X_{obs}$ and let $\tilde{\sigma}^2$ be a draw based on the deviations $(y_{obs} - X_{obs}\hat{\beta})'(y_{obs} - X_{obs}\hat{\beta})/\tilde{g}$, where $\tilde{g}$ is a draw from a $\chi^2$ distribution.

2. Draw $\tilde{\beta} = \hat{\beta} + \tilde{\sigma} z_1 V^{1/2}$, where $V^{1/2}$ is the square root of the Cholesky decomposition of the cross-products matrix $S = X_{obs}'X_{obs}$, and $z_1$ is $p$-dimensional vector of $N(0,1)$ random variates.

3. Calculate $\tilde{\eta}(i,j) = |X_{obs,[i]}\hat{\beta} - X_{miss,[j]}\tilde{\beta}|$; $i = 1, 2, \ldots, n_1$, $j = 1, 2, \ldots, n_0$.

4. Construct $n_0$ sets $W_j$ containing $d$ candidate donors from $y_{obs}$ such that $\sum_d \tilde{\eta}(i,j)$ is minimum. Break ties randomly.

5. Randomly draw one donor $i_j$ from $W_j$ for $j = 1, 2, \ldots n_0$.

6. Impute $\tilde{y}_j = y_{i_j}$, for $j = 1, 2, \ldots, n_0$. 

11 / 53
Bayesian Regression Imputation: mice.impute.norm()

- Bayesian imputation under the normal model proceeds much like predictive mean matching

1. Obtain $\hat{\beta}$ based on $X_{obs}$ and let $\tilde{\sigma}^2$ be a draw based on the deviations $(y_{obs} - X_{obs}\hat{\beta})' (y_{obs} - X_{obs}\hat{\beta})/\tilde{g}$, where $\tilde{g}$ is a draw from a $\chi^2$ distribution.

2. Draw $\tilde{\beta} = \hat{\beta} + \tilde{\sigma} \tilde{z}_1 V^{1/2}$ as before.

3. Calculate the imputed value $\tilde{y}$ as $\tilde{y} = X_{miss}\tilde{\beta} + \tilde{z}_2 \tilde{\sigma}$, where $\tilde{z}_2$ is a $j = 1, 2, \ldots, n_0$ vector of $N(0, 1)$ random variates for those with missing data on $y$.

- A new $\tilde{y}$ is obtained by drawing a new $\tilde{\sigma}^2$. This can be repeated $m$ times.
For PISA, most of the variables in the CQ are ordered categorical (e.g. Likert scales).

We wish to take the correct probability model into account and so for ordered categorical variables we use the proportional odds logistic regression model.

Approach similar to Bayesian imputation except

1. Obtain $\hat{\beta}$ by iteratively reweighted least squares

2. Obtain $\tilde{p} = 1/(1 + \exp(-X_{\text{miss}}\hat{\beta}))$

3. Use the cumulative logit to assign $K$ categorical responses

$$\log \left[ \frac{p(\tilde{y}_j \leq k)}{p(\tilde{y}_j > k)} \right] = \log \left[ \frac{\tilde{p}_{1j} + \ldots + \tilde{p}_{kj}}{\tilde{p}_{1,k+1} + \ldots + \tilde{p}_{Kj}} \right]$$
Some Preliminary Results

- We study predictive mean matching, Bayesian linear regression, and proportional odds logistic regression using data from the US sample of PISA 2012.

- We impute all missing data in the CQ. This includes missing by design as well as item missing data. We do not use the PVs in the imputation.

- Missing data are assumed to be MCAR or MAR.

- We request 1 imputation and 5 imputations for each procedure.

- We use the R software program “mice” (van Buuren & Groothuis-Oudshoorn, 2011).
Figure 1: PMM for selected rotated CQ items with 1 imputation. ST48Q01 (Maths Intentions: Form A); ST37Q01 (Math self efficacy: Form B); ST57Q01 (Out of school time: Form A); ST80Q01 (Cognitive activation: Form C)
Figure 2: PMM for selected rotated CQ items with 5 imputations.
Figure 3: POLR and NORM for selected rotated CQ items with 1 imputation.
Figure 4: POLR and NORM for selected rotated CQ items with 5 imputations. Each booklet imputed separately. Cognitive activation imputed with fewer items in the booklet.
For any rigorous study of imputing CQ data in a rotation design, validity criteria must also be established.

Here we use the work of Rässler (2002) and her colleagues.

1. First Level Validity: Preserving individual values
2. Second Level Validity: Preserving joint distributions
3. Third Level Validity: Preserving correlation/covariance structure

We adapt these levels to the questionnaire rotation design problem.
Let $x$ be an observed item in Form A and let $y$ be an observed item in Form B.

Let $\tilde{x}$ be the imputed value of $x$ in Form B and let $\tilde{y}$ be the imputed value of $y$ in Form A.

Let Form $\tilde{A}$ be the imputed Form A and Form $\tilde{B}$ be the imputed Form B.

Under the assumption of a randomized questionnaire rotation, we expect

1. $f_x = f_{\tilde{x}}$
2. $cov(x, \tilde{y}) = cov(\tilde{x}, y)$
3. $R^2_{PV.\tilde{A}} = R^2_{PV.\tilde{B}} = R^2_{PV.\tilde{C}}$
### Table 1: Correlations of selected observed and imputed items

<table>
<thead>
<tr>
<th>Form</th>
<th>Variable 1</th>
<th>Variable 2</th>
<th>Correlation</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 1$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>ST37Q01</td>
<td>ST57Q01</td>
<td>-0.205</td>
</tr>
<tr>
<td>Ĉ</td>
<td>ST37Q01</td>
<td>[ST57Q01]</td>
<td>-0.140</td>
</tr>
<tr>
<td>B</td>
<td>ST80Q01</td>
<td>ST37Q01</td>
<td>0.151</td>
</tr>
<tr>
<td>Ā</td>
<td>[ST80Q01]</td>
<td>ST37Q01</td>
<td>0.091</td>
</tr>
<tr>
<td>$m = 5$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>ST37Q01</td>
<td>ST57Q01</td>
<td>-0.205</td>
</tr>
<tr>
<td>Ĉ</td>
<td>ST37Q01</td>
<td>[ST57Q01]</td>
<td>-0.127</td>
</tr>
<tr>
<td>B</td>
<td>ST80Q01</td>
<td>ST37Q01</td>
<td>0.151</td>
</tr>
<tr>
<td>Ā</td>
<td>[ST80Q01]</td>
<td>ST37Q01</td>
<td>0.085</td>
</tr>
</tbody>
</table>
Table 2: $R^2$ from regression of math $PV1$ on all items in each form.

<table>
<thead>
<tr>
<th></th>
<th>Form A</th>
<th>Form B</th>
<th>Form C</th>
</tr>
</thead>
<tbody>
<tr>
<td>$m = 1$</td>
<td>0.820</td>
<td>0.816</td>
<td>0.836</td>
</tr>
<tr>
<td>$m = 5$</td>
<td>0.820</td>
<td>0.812</td>
<td>0.832</td>
</tr>
</tbody>
</table>


**Table 3:** $R^2$ from regression of math $PV1$ on all items in each form by lowest and highest 25% of the index of economic social and cultural status (ESCS).

<table>
<thead>
<tr>
<th>Subgroup</th>
<th>Form $\tilde{A}$</th>
<th>Form $\tilde{B}$</th>
<th>Form $\tilde{C}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$R^2$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower 25% ESCS</td>
<td>0.850</td>
<td>0.832</td>
<td>0.878</td>
</tr>
<tr>
<td>Upper 25% ESCS</td>
<td>0.854</td>
<td>0.868</td>
<td>0.845</td>
</tr>
<tr>
<td>$R^2_{adj}$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lower 25% ESCS</td>
<td>0.550</td>
<td>0.538</td>
<td>0.606</td>
</tr>
<tr>
<td>Upper 25% ESCS</td>
<td>0.541</td>
<td>0.568</td>
<td>0.548</td>
</tr>
</tbody>
</table>

*Note.* Results based on predicted mean matching with one imputation. The DV is the first plausible value and all predictors for each form are used.
The pros and cons of CQ rotation notwithstanding, it is essential that parallel studies of alternative imputation methods relevant to CQ rotation be conducted.

For this simple case study we found that predictive mean matching with 1 or 5 imputations performed very well, even without the use of the PVs as part of the imputation.

Using PVs didn’t change the results.

The approach of combining Bayesian regression and proportional odds logistic regression worked less well and deserves careful study.

Other methods require additional study.
We also argued that validity studies must be conducted.

We provided an approach to assessing the validity of the imputations that is in the spirit of Rässler’s (2002) approach and reasonable for the sampling design and CQ rotation design.

Preliminary evidence from the regressions suggests that Forms $\tilde{A}$, $\tilde{B}$ and $\tilde{C}$ might be “exchangeable” giving rise to the same relations to the PVs.

Different methods of CQ imputation followed by PV estimation remains to be studied.

We are currently engaged in this work and hope our results will be used to inform the CQ design for PISA and other ILSAs.
Part II. Data Fusion of Large Scale Assessments

- The OECD Program for International Student Assessment (PISA) and the Teaching and Learning International Survey (TALIS) constitute two of the largest ongoing international student and teacher surveys presently underway.

- Data generated from these surveys offer researchers and policymakers opportunities to identify particular educational institutional arrangements – that is, how aspects of educational systems are organized to promote equality of educational opportunity both within and between countries.

- Naturally, policy makers are interested in all three levels of the school system – students, teachers, and schools, in order to fully understand within and between country differences in relations between the inputs, processes, and outcomes of education.
Each survey is missing an important component of the educational system in their design – namely, PISA is missing teacher level data and TALIS is missing student level data.

Because a simultaneous administration of both surveys may not be feasible for many countries, this limits the extent to which information unique to each survey can be understood jointly.

A more feasible approach to linking the PISA survey to the TALIS survey involves the creation of a synthetic cohort of data - that is, a new data file that combines information from both surveys.

Two approaches are common and will be explored in this study. 

1. Statistical matching

2. Imputation

We will refer to both approaches in this context as data fusion.
We evaluate the extent to which each method provides a synthetic data set that maintains the essential properties of PISA and TALIS, concentrating on a set of validity criteria established by Rässler (2002).

Our evaluation relies on an experimental comparison of the validity of each method relative to a standard. For this purpose, we use data from Iceland.

We chose Iceland because it is the only OECD country that implemented PISA and TALIS on the population of PISA students, all TALIS teachers, and all PISA and TALIS schools.
Although missing-completely-at-random (MCAR) is usually not realistic in applied settings there is one unique situation in which MCAR might be reasonably assumed to hold – and that is where the missing data are missing by design.

In the case of PISA and TALIS, the two surveys have no units in common but do have variables in common – in particular, variables from the survey of principals in both the PISA and TALIS samples.

Because there are no units in common across the two surveys, the missing data are reasonably considered to be MCAR.
Validating Imputation Methods

Here too, for any rigorous study of data fusion, validity criteria must also be established.

Once again, we use the work of Rässler (2002) and her colleagues.

1. First Level Validity: Preserving individual values
2. Second Level Validity: Preserving joint distributions
3. Third Level Validity: Preserving correlation/covariance structure
   - The conditional covariance of variables not jointly observed (conditioning on the common variables) should be zero.
In the interest of time, we examine only Bayesian bootstrap predictive mean matching and the EM-bootstrap. In our paper, we also examined:

1. Nonparametric Hot Deck Matching: We use the R program *StatMatch*

2. Stochastic regression imputation: We use *mice*

3. Bayesian linear regression via chained equations: We use *mice*
The goal of BBPMM is to further relax the distribution assumptions associated with draws from the posterior distributions of the model parameters.

The algorithm begins by forming a Bayesian bootstrap of the observations (Rubin, 1981).

The Bayesian bootstrap (BB) is quite similar to conventional frequentist bootstrap (Efron, 1979).

It provides a method for simulating the posterior distribution of the parameters of interest rather than the sampling distribution of parameters of interest.

It is more robust to violations of distributional assumptions associated with the posterior distribution.
Next, BBPMM obtains estimates of the regression parameters from the BB sample. This is followed by the calculation of predicted values of the observed and missing data based on the regression parameters from the BB sample.

Predictive mean matching is then performed and these steps can be carried out \( m \geq 1 \) times to create \( m \) multiply imputed data sets.

We use the R software program \( BaBooN \).
A Hybrid Method: The EM Bootstrap

- The first step is to bootstrap the PISA and TALIS concatenated data set to create \( m \) versions of the incomplete data, where \( m \) ranges typically from 3 to 5 as in other multiple imputation approaches.

- Bootstrap resampling involves taking a sample of size \( n \) with replacement from the original dataset.

- Here, the \( m \) bootstrap samples of size \( n \) are obtained from the PISA and TALIS concatenated file, where \( n \) is the total sample size of the file.

- Second, for each bootstrapped data set, the EM algorithm is run but priors could be added.
Because $m$ boostrapped samples are obtained, and that each EM run on these samples may contain priors, then once the EM algorithm has run, the model parameters will be different.

With priors, the final results are the *maximum a posteriori* (MAP) estimates; the Bayesian counterpart of the maximum likelihood estimates.

Missing values are imputed based on the final converged estimates for each of the $m$ datasets.

These $m$ versions can then be used in subsequent analyses.

We use the R software program *Amelia* for the EM algorithm.
142 schools participated in either the TALIS survey or the PISA survey.

Of these, 122 PISA and TALIS schools were able to be matched.

The 20 schools that were unmatched were eligible for TALIS or PISA, but not both.

An additional 39 schools were excluded due to large amounts of missing data on variables needed for the fusion procedures. Finally, 5 schools were excluded because they were identified to be influential outliers.

Thus, the statistical matching procedures utilize data from 78 schools in Iceland with full information from the PISA and TALIS data sets.
Preliminary analyses indicated that randomly deleting data would yield a sample size that was likely too small to effectively judge the quality of the matching procedures.

We duplicated the Iceland data and then removed PISA data for half the sample and TALIS data for the other half of the sample.

This led to a sample of 78 schools with PISA data and 78 schools with TALIS data.

Because the duplication and subsequent deletion of the data were not dependent on any of the observed PISA, TALIS or common variables, the missing data are missing completely at random.
Matching Variables

- School sector;
- The size of the school community; the total enrollment in the school;
- A measure of the availability of school material resources;
- The extent to which teacher absenteeism interferes with student learning;
- A measure of the extent to which student-related factors affect the school climate;
- A measure of the disciplinary climate of the school.
Unique Variables

- From PISA:
  - Enjoyment of reading; summarizing skills. Both measures are averaged to the school level for analysis.

- From TALIS:
  - Teacher job satisfaction; teacher self-efficacy. These measures were also averaged to the school level for analysis.
Figure 5: Kernel Density and QQ Plots For Predictive Mean Matching
Table 4: Conditional Correlation Matrix For Original Iceland Data and Predictive Mean Matching

<table>
<thead>
<tr>
<th></th>
<th>y1</th>
<th>y2</th>
<th></th>
<th>y1</th>
<th>y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>-0.02</td>
<td>-0.04</td>
<td>x1</td>
<td>-0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>x2</td>
<td>0.01</td>
<td>-0.01</td>
<td>x2</td>
<td>0.00</td>
<td>-0.03</td>
</tr>
</tbody>
</table>
Figure 6: Kernel Density and QQ Plots For Bayesian Bootstrap Predictive Mean Matching
### Table 5: Conditional Correlation Matrix For Original Iceland Data and Bayesian Bootstrap Predictive Mean Matching

<table>
<thead>
<tr>
<th></th>
<th>y1</th>
<th>y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>-0.02</td>
<td>-0.04</td>
</tr>
<tr>
<td>x2</td>
<td>0.01</td>
<td>-0.01</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>y1</th>
<th>y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>0.01</td>
<td>0.01</td>
</tr>
<tr>
<td>x2</td>
<td>0.01</td>
<td>0.02</td>
</tr>
</tbody>
</table>
Figure 7: Kernel Density and QQ Plots For EM Bootstrap
Table 6: Conditional Correlation Matrix For Original Iceland Data and EM Bootstrap Without Priors

<table>
<thead>
<tr>
<th></th>
<th>y1</th>
<th>y2</th>
<th></th>
<th>y1</th>
<th>y2</th>
</tr>
</thead>
<tbody>
<tr>
<td>x1</td>
<td>-0.02</td>
<td>-0.04</td>
<td>x1</td>
<td>0.01</td>
<td>0.00</td>
</tr>
<tr>
<td>x2</td>
<td>0.01</td>
<td>-0.01</td>
<td>x2</td>
<td>-0.00</td>
<td>-0.02</td>
</tr>
</tbody>
</table>
Summary

- We find that Bayesian bootstrap predictive mean matching, and the EM-bootstrap worked quite well with respect to Rässler’s (2002) third and fourth level validity criteria.

- We anticipate that the implementation of priors would influence the comparability of the method to other methods depending on the precision of the priors.

- Findings from matching current cycles of PISA and TALIS could be used to inform the specification of priors for future statistical matching exercises. This is a topic of our current research.
Statistical matching is typically limited to single level data structures.

However, there does exist a two-level statistical matching algorithm in the R software program *mice* based on the Gibbs sampling algorithm.

To summarize, conducting a data fusion of PISA and TALIS might be a reasonable option for countries that are unable to administer both surveys to the same sample schools. And this will be particularly important for the upcoming PISA/TALIS link in 2018.
This talk focused on two major issues in the design and analysis of large scale assessments.

1. Rotation issues in the design of context questionnaires
2. Fusion of different data sets.

Both issues are based on the same underlying methods for handling missing data.

Both issues require the development and application of validation methods for missing data imputation.
Challenges with CQ Rotation

There are numerous challenges associated with CQ rotation that have yet to be addressed.

1. Separate rotation of CQ and cognitive booklets v. simultaneous rotation and imputation of both?
2. Operational issues and time constraints.
3. What type of data set will users have access to?
4. Small area estimation validity?
There are numerous opportunities available with CQ rotation:

1. The major benefit of CQ rotation is expanded content to place the literacy scores in context.

2. Expanded CQ rotation also permits broader coverage of non-cognitive outcomes (attitudes, aspirations, etc).

3. Greater opportunity for policy relevant modeling.
Numerous challenges face the use of data fusion.

1. Multilevel data fusion?
2. Sampling weights – what is the relevant population?
3. Temporal concerns? When were surveys obtained?
Numerous opportunities await more data fusion.

1. Cost savings?

2. Greater coordination and collaboration in national data collection efforts?

3. Potential for great policy relevance - particularly when data are fused with national system indicators.
GRAZIE MILLE