**PAPER NUMBER #105**

**Push Button Replication: Is impact evaluation evidence for international development verifiable?**

**Presenting Authors**

Dr Benjamin Wood1, Dr Annette Brown2

**Affiliation**

1. International Initiative for Impact Evaluation
2. FHI 360

**Country of residence**

United States of America

**Objectives/aims**

Our paper presents the results of the 3ie’s “push button replication” (PBR) project. The project establishes PBR procedures and standards, to better align authors’ and PBR researchers’ expectations around PBR verification. It also tests whether impact evaluations of development programs in low- and middle-income countries are generally push button verifiable.

PBR tests the validity of published results using the original study’s data and programming code. The premise is that third party researchers should not need to write new code to reproduce published results. PBR asks the question can we use the paper’s data and the original authors’ programming code to reproduce the published results.

**Methods**

We test the push button replicability of 109 development impact evaluations. Our sample includes all impact evaluations published in 2014 in the top ten journals for these studies. We determined the top ten journals as those that published the greatest number of development impact evaluations between 2010 through 2012, as catalogued in 3ie’s impact evaluation repository. These studies span several disciplines including public health, political science and economics. Each PBR researcher followed 3ie’s PBR protocol, with their PBR reports posted on the Open Science Framework.

**Main findings**

Of the 109 published articles in our sample, 27 are fully push button replicable. Of those 82 that are not, 15 did not provide adequate replication materials and 10 include results that significantly differed from the original results. Of the 57 studies for which authors would not provide data or documentation, 7 include text in their papers stating the data are available upon request. The good news is that 14 of the 109 articles in our sample had open data. Our project revealed a wide variation in the practices of researchers for maintaining data and documentation. We also found a significant difference in the PBR results between articles published in social science public health journals.