



BASIE Tool Framing and Instructions: Road Map

BASIE Tool Framing and Instructions: Road Map

Teen Pregnancy Prevention (TPP) researchers may want to use an alternate lens to guide interpretation of the effects of their programs beyond statistical significance, partly in response to the statement from the American Statistical Association in 2016 and follow-up commentaries in 2019 in *The American Statistician and Nature* (Wasserstein & Lazar, 2016; Wasserstein et al., 2019; Amrhein et al., 2019). The Bayesian Interpretation of Estimates (BASIE) is a framework for interpreting impact estimates that moves beyond statistical significance. Specifically, it offers researchers opportunities to make more intuitive probabilistic statements about the impacts of programs, such as “There is an 83 percent probability that Program X favorably affected treatment student outcomes” even if the impact estimate associated with that program was not statistically significant at conventional levels.

This document is intended to complement the introductory webinar on Bayesian interpretation and the BASIE tool offered to earlier TPP grant recipients conducting rigorous evaluations of innovative programs. The BASIE probability tool and webinar are available [here](#), along with an example version of the BASIE tool, with findings already entered into the Excel file. See Deke et al. (2022) for more details on BASIE.

The first three sections of this document are titled to align with the tab names in the Excel BASIE tool: (1) Prior evidence, (2) Study findings, and (3) Interpretation. A final section provides guidance on ways to use the Interpretation tab to present evaluation results – whether in final reports or journal articles.

Throughout this document, we highlight how to enter data into the BASIE tool and describe how comparable information was entered into the BASIE example available in the link above.

Prior evidence

To calculate the posterior probabilities of interest (that is, the probability that a given program had a favorable effect), it is necessary first to identify the distribution of prior evidence, which may, for example, include evidence from recent meta-analyses of impact evaluations. See Deke et al. (2022) and the BASIE webinar for a discussion of considerations for selecting prior evidence and for examples specific to TPP, respectively.

This first tab enables the user to enter key information about the evidence base. In the BASIE tool, the evidence base is documented as the “estimated prior distribution,” which represents the distribution of outcomes from which a given study’s findings might be drawn.

Example BASIE walkthrough

The tab in the example version of the BASIE tool includes four prior distributions. The prior distribution from Juras et al. (2019) represents the mean and standard deviation (SD) of the distribution of impacts from a meta-analysis of teen pregnancy prevention programs conducted since 2010 that examined distal (as opposed to proximal) sexual behavior outcomes. To use prior distributions for proximal outcomes, the tab also includes information from Herrmann et al. (2019), showing the mean and SD of the distribution of



impacts from a broad range of programs included in the What Works Clearinghouse (WWC) (a systematic review focused on education programs). Given that the WWC includes many researcher-developed education outcomes that are well aligned with the programs being tested, the Herrmann et al. (2019) prior distribution could be a better indicator of the effects of TPP interventions on the proximal outcomes of knowledge, attitudes, and intentions typically assessed in many TPP evaluations. Notably, the mean for Herrmann et al. (2019) is larger than the mean for Juras et al. (2019), aligning with expectations that impacts for proximal outcomes tend to be larger, on average, than impacts for more distal sexual behavior outcomes.

In addition, there are rows for these two sources of evidence (Juras et al. [2019] and Herrmann et al. [2019]) that are labeled “zero-centered” – that is, where the SDs have been maintained, but the means have been shifted to zero. Zero-centered distributions may be particularly useful when thinking about interpreting the effects of studies that compare two competing programs (rather than a comparison to a business-as-usual control group) (Deke, Finucane, and Thal, 2022).

Instructions for BASIE

Step 1 (Column A). Enter the source(s) of evidence that you plan to use for the mean and SD for the prior distributions. In most cases, the source will be a study citation. The purpose is to list your evidence sources and enter a clear identifier for each source so that all contributing sources of evidence may be easily referenced and accessed in subsequent tabs. Include a new source-of-evidence reference for each new mean and SD that you plan to use as prior evidence.

Step 2 (Column B). After you have entered the names of the source(s) of evidence, the next step is to enter the mean of the prior distribution from each source in Column B. To provide a common scale for the prior evidence information, the means should be presented in effect size units. In addition, all effect sizes should be coded so that a positive sign corresponds to a favorable effect. Estimated impacts on dichotomous outcomes (that is, whether a participant is sexually active or not) should also be converted into effect size units. The What Works Clearinghouse (2022) provides details on how to convert dichotomous outcomes into effect size units.

Step 3 (Column C). Enter the SD for each prior distribution associated with a given source of evidence.

BASIE tool calculation. After you enter the means and SDs for all the prior sources of evidence, the tab will automatically generate a distribution of prior program impacts in Columns D through I. Each column provides an estimated proportion of prior program impacts that is below (Columns D through F) or above (Columns G through I) a pre-established effect size threshold. In total, six columns correspond to specific thresholds. The information provides a rough sense of the prevalence of distributions of impacts of various sizes for a given prior distribution.

Study findings

The purpose of this tab is to document the impact estimates (reported in effect size units) and standard errors of each outcome estimated from your study. The information you enter in this tab will be pulled into the interpretation tab and combined with the information from the previous evidence tab to produce a



Bayesian posterior probability (the probability that the program had a favorable effect for study participants, given the information observed in the study and the information about prior evidence) and other statistics.

Example BASIE walkthrough

The tab includes four sets of impact estimates and their associated standard errors from two hypothetical studies: Study A and Study B. The rows of findings from both Study A and Study B report the means and standard errors of a program's estimated impacts on (1) STI knowledge and (2) sexual initiation. In this example, the impact estimates on STI knowledge and sexual initiation from both studies are 0.25 and 0.10, respectively. The notable difference in this example is that the standard errors associated with the estimates from Study A (0.10) are smaller than the standard errors associated with Study B (0.20). This difference is likely attributable to Study A's larger sample size compared to Study B's sample size. In addition, a note in the tab indicates that the program's negative and favorable effect on (reducing) sexual initiation was entered in the worksheet as a positive number to simplify interpretation. In many TPP programs, a reduction in certain behaviors is considered a favorable outcome, and entering such information as a positive impact in the BASIE tool will always make it easier to interpret positive numbers as favorable results.

Instructions for BASIE

Step 1 (Column A). Document the specific outcome measure (and outcome domain, if that will aid in interpretation) for each impact estimate. In addition, enter the appropriate study identifiers if you plan to calculate posterior probabilities from several studies in a single worksheet. Include a new row for each specific outcome measure in your study for which you want to generate a posterior distribution.

Step 2 (Column B). Enter the impact estimates from your study in Column B. As noted, the estimates should be expressed as an effect size to be compatible with the prior distribution. You should enter an effect size for every outcome for which you are estimating a posterior probability statistic.

Step 3 (Column C). Enter the standard error of each impact estimate. Along with the impact estimates that you entered in the previous step, the standard errors from your study will feed into the interpretation tab. You should enter a standard error for every outcome for which you are estimating a posterior probability statistic.

Interpretation

The interpretation tab brings together all the information entered in the previous tabs to produce the posterior probability and other statistics of interest. These statistics allow you to make statements about the probability that the effect of the program you are examining on a specific outcome is favorable (or larger/smaller than a given threshold in the distribution of evidence for comparable programs). The tab is divided into three color-coded sections that correspond to: (1) the prior distribution (purple); (2) your study findings, known as the likelihood (green); and (3) the posterior statistics of interest (blue). Below, we briefly review each of these sections to clarify the information and steps you need to follow to interpret the Bayesian results.



Example BASIE walk-through

The first section is color-coded purple and is a summary of the prior distribution(s) based on the information entered in the prior evidence tab. In this section, the effect size means and SDs from the Herrmann et al. (2019) and Juras et al. (2019) studies are listed for the two outcomes of interest. In addition, the means and SDs for the zero-centered priors are included. The second section is color-coded green and shows the outcomes, impact estimates, and standard errors for the impacts examined in Study A and Study B. This information is pulled automatically from the study findings tab. The third section is color-coded blue and provides information from the posterior distribution(s): the Bayesian impact estimates and posterior SDs and the probability statement(s) that the true effect size of the program(s) examined in Study A and Study B is greater or less than an effect size threshold determined to be substantively important. The BASIE tool automatically generates Bayesian impact estimates and posterior SDs based on the information from the earlier tabs.

Row 6, for example, summarizes the probability that the true effect of the program in Study A on STI knowledge was greater than zero. In this case, the “greater than” (Column I) option was selected, and the reference effect size (Column J) was chosen to be zero. Putting this all together, an appropriate interpretation of the output in row 6 follows:

“We estimate a 99 percent probability that the true effect of the program tested in Study A on STI knowledge was greater than zero, given our estimates and prior evidence on the impacts of education interventions.”

The same approach to interpretation should be used for the other rows in the tab. As noted, users can change the effect size thresholds in Column J by manually entering different values of interest. For example, rather than an effect size of zero, an effect size of 0.20 could be used. With this change, the posterior probability for row 6 would change from 0.99 to 0.67. In this case, the interpretation would change as follows:

“We estimate a 67 percent probability that the true effect of the program tested in Study A on increased STI knowledge was greater than 0.20, given our estimates and prior evidence on the impacts of education interventions.”

Note that each row in the interpretation tab has an associated posterior probability for each outcome examined.

Instructions for BASIE

Step 1 (Column A). Enter the source of evidence for your prior distribution. Column A includes a drop-down box that lets you select among the sources of evidence that you listed in the “prior evidence” tab.

Step 2 (Column D). Enter the outcome of interest from your study. Column D contains a drop-down box that lets you choose among the outcomes that you entered in the study findings tab.

Step 3 (Columns I and J). In Column I, select whether to assess if the Bayesian impact estimate is greater than (>) or less than (<) a defined effect size value. The column contains a drop-down that lets



you select the direction of the effect size comparison. Next, in Column J, enter a value for the reference effect size against which you want to compare the Bayesian effect size. The effect size value can be zero or any negative or positive value. Enter an effect size value that makes substantive sense given the program context or use zero to interpret the probability that the program’s effect is favorable/unfavorable.

BASIE tool calculation

After entering the information in the columns specified above, the tab will automatically populate the remaining columns. For example, when you enter the source of evidence for your prior distribution (Column A), the mean and SD for that prior distribution will automatically fill Columns B and C. Similarly, after entering the study outcome of interest in Column D, the impact estimate from your study and its standard error will automatically populate Columns E and F, respectively. Finally, after entering the direction of the effect size comparison in Column I and the effect size threshold in Column J, the calculated posterior probability (a value between 0 and 1) will appear in Column K. The probability shown in Column K is the posterior probability, which tells you the probability that the true effect from your study was greater or less than the reference effect size. This probability will appear for every outcome of interest that is entered into the tab, allowing you to compare posterior probabilities for outcomes under different assumptions about the prior distribution and at different effect size thresholds.

Presentation

An important final step in using the BASIE tool is extracting and condensing the most relevant information from the interpretation tab and summarizing the results in a table for a given audience. The BASIE tool does not currently have built-in functionality for generating tables, but the information summarized can be manually extracted and placed into a table for presentation and dissemination.

Example BASIE walk-through

Rows 6 and 14 of the interpretation tab show the results for two outcomes examined in Study A: STI knowledge and sexual initiation. For both outcomes, specific information from the study findings and posterior distribution sections of the interpretation tab can be extracted and pasted into a table to simplify the results and messages for some audiences. In Table 1, we present the impact estimates, standard errors, Bayesian impact estimates, and probabilities that the true effect of the program is greater than zero.

Table 1. Bayesian impact estimates tested by the program in Study A on STI knowledge and sexual initiation

| Impact estimate description | Impact estimate | Standard error | Bayesian impact estimate | Probability that the true effect > 0 |
|-------------------------------------------------------------|-----------------|----------------|--------------------------|--------------------------------------|
| Impact on STI knowledge (risk and protective factor domain) | 0.25* | 0.10 | 0.24 | 99% |
| Impact on sexual initiation (sexual behavior domain) | 0.10 | 0.10 | 0.08 | 83% |

Risk and protective factor used Herrmann et al. (2019) as prior evidence; sexual behavior used Juras et al. (2019) as prior evidence.



References

- Agodini, R., Harris, B., Seftor, N., Remillard, J., & Thomas, M. (2013). *After two years, three elementary math curricula outperform a fourth*. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Amrhein, V., Greenland, S., & McShane, B.B. (2019). Scientists rise up against statistical significance. *Nature*, (567), 305–307.
- Deke, J., Finucane, M., & Thal, D. (2022). *The BASIE (BAYesian Interpretation of Estimates) framework for interpreting findings from impact evaluations: A practical guide for education researchers*. NCEE 2022-005. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. [The BASIE \(BAYesian Interpretation of Estimates\) framework for interpreting findings from impact evaluations: A practical guide for education researchers](#)
- Juras, R., Tanner-Smith, E., Kelsey, M., Lipsey, M., & Layzer, J. (2019). Adolescent pregnancy prevention: Meta-analysis of federally funded program evaluations. *American Journal of Public Health*, 109(4), e1–e8.
- Herrmann, M., Clark, M., James-Burdumy, S., Tuttle, C., Kautz, T., Knechtel, V., Dotter, D., Wulsin, C. S., & Deke, J. (2019). *The effects of a principal professional development program focused on instructional leadership*. NCEE 2020-0002. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance.
- Wasserstein, R. L., Schirm, A. L., & Lazar, N. A. (2019). Moving to a world beyond “ $p < 0.05$.” *The American Statistician*, (73, sup. 1), 1–19. DOI: [10.1080/00031305.2019.1583913](https://doi.org/10.1080/00031305.2019.1583913)
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA’s statement on p -values: Context, process, and purpose. *The American Statistician*, 70(2), 129–133.
- What Works Clearinghouse. (2022). *What Works Clearinghouse procedures and standards handbook*, Version 5.0. U.S. Department of Education, Institute of Education Sciences, National Center for Education Evaluation and Regional Assistance. <https://ies.ed.gov/ncee/wwc/Handbooks>.