# Brady Education Foundation

**Stage 1 Application**
**Proposal Review Criteria**
*Updated August 19, 2020*

**DISQUALIFIED APPLICATIONS:**
**Proposals that meet any of the following disqualifying criteria will <u>NOT</u> be reviewed by the Full Board of BEF:**

- Mission
    - Specific aims are outside the mission of the Foundation (posted on website) (Section 2 of application: questions 1 and 2)
        - e.g., target population is not defined by race or income
        - e.g., ultimate intended outcomes of the program to be evaluated do not include promoting academic, achievement or cognitive performance or growth
- Funding policies (posted on website)
    - Proposed work concerns communities of color, but the leadership of the research team (PI/co-PI level) does not include any researchers of color (Section 5: question 3)
    - Project is outside the US
    - Evaluation agency is a for-profit company
    - Program to be evaluated is for-profit
    - PI not a researcher (e.g., is a development officer)
    - Proposal is not run through the PI's home institution
    - Budget (Section 6 and budget justification)
        - Budget is inconsistent with Indirect Costs policy
            - Indirect costs included as direct costs
            - Indirect costs are included when total budget is $50,000 or less
            - For grants with total costs higher than $50,000, indirect costs included are higher than 10% for primary institution or one or more subcontract
            - The primary institution charges indirect costs on one or more subcontracts
        - One or more salary is above the salary cap (see Budget instructions in application)
        - Conference support request is inconsistent with travel policy (see Budget instructions in application)
        - Budget includes funding for scholarships, capital expenses, scaling up program, or continuing education (Note: GRA tuition is allowed when appropriately justified)
- Application
    - Missing responses to questions in the application (Sections 1-6)
    - Proposed project dates
        - Not within the range for the application cycle (see dates on website)
            - Rare exceptions can be made when a compelling case is articulated; rationale to be provided in cover letter and in application (Section 5: question 4)
        - Duration of funding requested is longer than 3 years
    - Budget
        - Missing budget information (e.g., salaries, FTEs; subcontracts) (Section 6)
        - Missing or incomplete budget justification (Section 7: required attachment)
    - Project timeline with benchmarks (Section 7: required attachment)
        - Missing or incomplete
        - Is over 3-page limit

- o CVs / resumes of key personnel (Section 7: required attachment)
  - Missing for one or more key personnel
  - One or more is over the 3-page limit
  - One or more does not include information requested (see instructions in application)
- o Letters of support from partnering organizations (Section 7: required attachment)
  - Missing for one or more partnering organizations (e.g., schools that would participate; data agreements when obtaining administrative data)
  - One or more does not include information requested (see instructions in application)
- o Approval letter from applicant institution / organization (Section 7: required attachment)
  - Missing for primary institution
  - Does not include information requested (see instructions in application)
- o Proof of tax-exempt status (Section 7: required attachment)
  - Missing for primary institution
  - Missing for one or more partnering institutions

_____

**APPLICATIONS REVIEWED BY THE FULL BOARD:**
**General Review Information**
Each specific review criterion (see below) is rated on a 5-point scale; higher scores indicate higher criteria strength. Reviewers consider strengths and weaknesses within each criterion when scoring.

| score | descriptor |
|-------|------------|
| 1 | Poor |
| 2 | Fair |
| 3 | Good |
| 4 | Excellent |
| 5 | Outstanding |

A reviewer's final recommendation concerning whether to invite or fund a proposal (depending on the Stage level) is not necessarily based on the average of the criterion scores. For example, a reviewer may give only moderate scores to some of the review criteria but still recommend inviting/funding because one criterion critically important to the project is rated highly; or a reviewer could give mostly high criterion ratings but still not recommend inviting/funding because one criterion critically important to the project being proposed is not highly rated.

In general, four overall criteria categories are considered when reviewing program evaluation applications (see below for more detail concerning each category)

1. Overall fit and potential impact
2. The program to be evaluated
3. The proposed research
4. Budget and budget justification

**Review Criteria Categories**

- Overall fit and potential impact
    - Aims and fit with BEF mission
    - Rationale (need for the program and need for this particular evaluation)
    - Collaboration
    - Potential impact and dissemination plans
- Program to be evaluated
    - Program implementation and goals
    - Target population
    - Feasibility of implementation from practitioner / service provider perspective
    - Accessibility from perspective of potential participants / target population
    - Affordability and Sustainability
    - Strength-Based orientation
- Research / Evaluation Criteria
    - Research / Evaluation design and methodology
        - Design
        - Intended study sample (size, recruitment, maintenance, demographics)
        - Procedures and measures
        - Analytic plan
        - Feasibility
    - Research / Evaluation environment and team
        - Research environment
        - Expertise of the PI / research team
        - Racial/ethnic composition of the leadership of the research team
- Budget and budget justification

**Detailed Information Concerning Review Criteria**
Below, the factors considered when scoring each of the criteria are described. The questions in the application most pertinent to each criterion are provided; however, responses throughout the application may be considered when scoring each criterion as well.

<u>Overall fit and potential impact</u>
The Brady Education Foundation seeks to close the educational opportunity gaps associated with race, ethnicity and family income. The Foundation pursues its mission by promoting collaboration among researchers, educators, and other stakeholders via the funding of research and program evaluations that have the potential of informing private funders and public policy. When determining the overall fit and potential impact of a proposal, the following criteria are considered:

- *Aims and fit with BEF mission*
    - Factors considered when scoring these criteria include what is/are the specific aim(s) of the proposed project (including primary aim of evaluating effectiveness as well as any secondary aims concerning possible mediators, moderators and/or cost-benefit analyses), clarity of the aim(s), and the goodness of fit of the aim(s) with the mission of the Foundation. Proposals that have clear aims that will address closing the educational opportunity gaps for children (ages birth through 18) associated with race and income as the primary aim, and those that seek to test moderators,

mediators and/or conduct cost-benefit analyses as secondary aims, score higher on these criteria. The Foundation tends not to fund studies that do not investigate main effects of a program as the primary aim (e.g., a proposal for which the primary aim is to investigating mediators of possible effects). (Section 2: questions 1 and 2; Section 1: questions 7 and 8 also considered if other current or potential funding partners exists and /or whether the proposed evaluation is part of a larger project)

- *Rationale (need for the program and need for this particular evaluation)*
  - o Factors considered when scoring this criterion include the extent to which compelling rationale is provided to justify the aim(s) of the proposed project, clarity of the overall Theory of Change guiding the proposed project and how *this specific project* fits within it, and the strength of the empirical literature supporting the aim(s) of the proposed project. Proposals that demonstrate a clear need for this program to help close the educational opportunity gaps, provide evidence supporting the components of the program to be evaluated, and demonstrate the clear need for the specific evaluation being proposed score higher on this criterion. Also considered is what "next steps" results from the proposed project might inform (e.g., providing pilot data that might lead to further program development and/or larger evaluation study). (Section 3: questions 1, 2 and 3)

- *Collaboration*
  - o Scoring of this criterion is based on the extent to which the program to be evaluated was originally developed via a prior collaboration between researchers and practitioners, and the extent to which the proposed project reflects strong collaboration between the research / evaluation team and practitioners / service providers (and other community stakeholders as appropriate such as parents / families, economists, policy makers, other community members, etc.) throughout the proposed project (e.g., developing questions, recruitment, data collection, analyses and interpretation, dissemination). Proposals that demonstrate authentic collaboration among the researchers, practitioners and other stakeholders score higher on this criterion. (Section 2: question 3)

- *Impact potential and dissemination plans*:
  - o Factors considered when scoring these criterion include potential dissemination products (e.g., solid contribution to lists such as the "What Works Clearing House" at the US Department of Education's Institute of Education Sciences, likelihood of publications in quality journals), and the potential impact for the results from the project to inform programmatic funding decisions by private foundations and/or public policy. Proposals that have specific plans for how results might be disseminated to researchers, practitioners, and policy makers, and those that are able to identify funders (e.g., specific foundations) and/or policy makers (e.g., public school districts) with whom they are in contact and whose work will be informed by the findings tend to score higher on these criteria. Proposals that have the potential to generalize beyond the specific program to be evaluated and inform how it might be scaled up to serve other communities tend to score higher as well. (Section 2: questions 4 and 5)

Program to be evaluated

Several factors concerning the program itself are considered when assessing whether to invite an application to Stage 2.

- *Program implementation and goals*
  - o Scoring of this criterion is based on the extent to which the program to be evaluated is likely to make a meaningful impact on closing the opportunity and resulting achievement gaps associated with race and income. Factors considered include the intended goals of the program and the activities participants engage in to achieve these goals, the duration / intensity of the program, and the methods the program uses to reach and enroll potential program participants. Proposals that provide compelling evidence that via the program outreach methods, the duration and intensity of the program, and the activities program participants engage in will likely have meaningful impacts on ultimately promoting academic, achievement, and cognitive outcomes score higher on these criteria. (Section 3: questions 1 and 4)

- *Target population*
  - o Factors considered when scoring this criterion include the extent to which the program to be evaluated is focused on minoritized racial / ethnic groups and low-income families. Programs that serve high proportions of children and families from these communities score higher on this criterion. (Section 3: question 9)

- *Feasibility of implementation from practitioner / service provider perspective:*
  - o Scoring of this criterion is based on the strength of the evidence provided concerning how well the program to be evaluated has been / can be implemented in the "real world", considering the opportunities and challenges faced by practitioners and service providers serving the target populations of the program, and the likelihood it could be scaled up across a variety of settings / communities. This criterion specifically concerns factors that facilitate the ability of those providing the program to implement it (factors that facilitate program participation of individuals in the target population are considered in the next criterion concerning accessibility). Proposals the identify specific challenges those providing the program potentially face implementating the program and how those are / might be addressed score higher on this criterion. (Section 3: question 5)

- *Accessibility from perspective of potential participants / target population*
  - o Scoring of this criterion is based on the strength of the evidence provided concerning how well the target population can access the program, considering both barriers to and incentives for program participation (such as the extent to which potential participants view the program as useful), particularly for individuals and families who typically face many barriers accessing high quality programs due to race or income. Proposals that provide data concerning the percentage of the target population that actually participate (both starting the program as well as continuing through its completion) and identify specific potential barriers to access and how those are / might be addressed score higher on this criterion. (Section 3: question 6)

- *Affordability and Sustainability*
  - o Affordability refers to the both the start-up costs required to initiate a program as well as the on-going operational costs of a program, weighing costs to potential benefits. Sustainability in this regard refers to the extent to which sources of funding sufficient to meet those costs can be identified (i.e., "financial stakeholders"). Scoring of this criterion is based on the strength of the evidence provided concerning the affordability and sustainability of the program to be evaluated. Proposals that provide estimates for both start-up and on-going implementation costs and specify

how those costs might be met by a community implementing the program (e.g., philanthropic donations; local, state, or federal funding) tend to score higher on this criterion.  (Section1: question 9; Section 3: question 7)

- *Strength-Based Orientation*
  - o Scoring of this criterion is based on the extent to which the program to be evaluated is grounded in a strength-based approach (rather than a deficit-based model), identifying specific strengths individuals and families bring that the program builds upon (i.e., while the program intends to build strengths by participating in the program, the program also recognizes strengths that participants bring that can be built upon to promote outcomes). Factors considered when scoring this criterion include the extent to which the program considers the specific and unique needs, challenges and strengths of children from diverse racial and socioeconomic communities, while also recognizing the variability that exists within populations in terms of needs, challenges and strengths (recognizing that one-size does not necessarily fit all). When the target population includes communities of color, proposals that identify specific strengths that develop within these communities that the program capitalizes upon to promote outcomes score higher on this criterion.  (Section1: question 9; Section 3: question 8)

Research Criteria

Two categories of research criteria are considered when assessing whether to invite an application to Stage 2: research design / methodology and the research environment / team.

- *Research / Evaluation design and methodology:* Scoring of these criteria are based on the extent to which the proposed design and methodology are likely to provide high quality data that will inform the specific aim(s) of the proposal. (Section 4: questions 1-9; Section 7: timeline with benchmarks)

  - o *Design*
    - ▪ Scoring for this criterion is based both on the clarity and the rigor of the evaluation design. In general, RCTs are considered more rigorous than comparison group designs, and comparison group designs are considered more rigorous that pre-post designs that only include program participants. Thus, RCTs tend to score higher than QED studies on this criterion, with both scoring higher than pre-post designs. Further:

      - For RCTs, scoring is also based on the clarity of the descriptions of the treatment and control groups (e.g., activities) and the clarity of the description of how and when randomization will take place. (Note: The Foundation does not favor withholding services in order to conduct an RCT, but does recognize that RCTs can be ethically done in a number of circumstances such as when programs are oversubscribed and thus can not serve all individuals who would want to participate, or when wait-list RCTs are conducted such that the control group is able to participate in the program after post data are collected).

      - For comparison group designs, scoring is also based on the strength of the rationale for not conducting an RCT and the extent to which possible confounding variables (e.g., due to selection bias) are identified and can be controlled for and baseline

equivalence among the groups can be determined. When methods such as propensity score matching are proposed, projects that demonstrate the ability to match on not only demographic factors score higher on this criterion (e.g., also match on performance measures prior to program participation).

- For pre-post designs that only include program participants, scoring is also based on the strength of the rationale for conducting neither an RCT nor a comparison group design and the rigor of the methods proposed to determine program effectiveness. (note: The Foundation very rarely funds proposals that propose pre-post designs)

- Secondary aims:
  Proposals that include secondary aims of investigating possible mediators and/or moderators of effects and/or conduct cost/benefit analyses tend to score higher than those only investigating main effects of the program. Scoring is based on design / methodology as well as the extent to which compelling rationale for including these secondary aims is provided (Section 3: question 1). Further:

  - Mediators:
    - Proposals that articulate clear hypotheses concerning what factors may account for the program effects and are able to operationalize them such the data collected can be entered into quantitative data analyses to test these hypotheses (e.g., path modeling) tend to score higher.

  - Moderators:
    - Such proposals may articulate clear hypotheses concerning how the strength or direction of a program's effects may vary for different subgroups (e.g., different racial groups, gender, dosage of program) or may aim to test whether the program is equally effective across racial, cultural and economic groups. Proposals that provide evidence of sufficient power to test these hypotheses and provide specific analytic plans for testing them (e.g., interaction terms entered into regression models) score higher.

  - Cost / benefits:
    - Proposals that are able to identify and quantify all costs of the program (both start-up and ongoing) as well as monetary benefits (could be at the individual or societal levels) and articulate a clear analytic plan score higher.

- *Intended study sample (size, recruitment, maintenance, demographics)*
  - Scoring for this criterion is based on the extent to which the sample size is judged to be appropriate for all aims of the proposed project including both main effect and moderation research questions (if applicable) and the extent to which the intended demographic characteristics are consistent with the aims of the study. Sample sizes should be supported by power analyses where appropriate (stating assumptions made). Also considered is whether recruitment procedures (and retention over-time if appropriate) are likely to recruit (and maintain) a sample representative of the intended population in sufficient

numbers to conduct the proposed project. Proposals that provide evidence that the target sample size results in the power needed to detect main effects (and test moderation hypotheses if applicable) and articulate specific strategies for recruiting (and maintain over time if applicable) a sample with the needed demographic characteristics to investigate the aim(s) of the study score higher. Studies that intend to recruit high proportions of children and families in minoritized racial / ethnic groups and low-income communities score higher on this criterion as well. Further:

- For RCTs, also considered when scoring this criterion are the sample sizes for each treatment and control group at the level of randomization and the strength of the evidence provided that the sample size will be large enough to detect meaningful effects (e.g., if randomizing at the classroom level, evidence that the number of classrooms included in each group is sufficient to detect effects).

- For comparison group designs, also considered when scoring this criterion are the strength of the evidence provided that the sample size will be large enough to detect meaningful effects and the strength of the evidence provided that the sample size is large enough to control for potential confounding variables.

- *Procedures and measures*
    - Scoring for these criteria are based on whether the proposed procedures and data to be collected are consistent with the aims of the proposed project and the extent to which outcome measures are consistent with the goals of the program (e.g., including procedures to assess student achievement if the ultimate goal of the program is to increase student achievement). Considered are the purpose of the data to be collected (e.g., to assess child outcomes), the methods used to collect data (e.g., direct assessment, parent-report, administrative data), and the strength of the evidence provided concerning the psychometric properties of the methods proposed. If sources for data include administrative data sets or other existing data sets currently managed by other parties, be sure to attach the data sharing agreements to the application. (Section 7, required attachment) Proposals that intend to obtain data from multiple sources, obtain child outcome data by directly assessing children, and those that demonstrate that the measures proposed are reliable and valid for the populations represented in the sample score higher on these criteria.

    - Note that the Foundation very rarely funds qualitative data collection and analyses (e.g., open-ended interviews, focus groups). The Foundation recognizes that qualitative data can provide very valuable information, but due to the Foundation's relatively small size and limited capacity, it is typically unable to funds this type of work. However, the Foundation does encourage applicants to seek other sources of funding for qualitative work when the Foundation is funding the quantitative data collection and analyses, recognizing that mixed-methods provide more complete understanding than either method alone.

- *Analytic plan*
    - Scoring for this criterion is based on the level to which the analyses for each specific aim of the proposed project are clearly articulated and appropriate (e.g., accounting for clusters in

modeling). Proposals that specify the statistical procedures that will be used to test the hypotheses (e.g., regression analyses, growth curve modeling, path analyses), explain them using the variables that will be included and in terms that educated individuals not familiar with the analytic methods can understand, and articulate the rationale for selecting the analytic methods (providing references as appropriate) score higher on this criterion.

- o *Feasibility*
    - Scoring of this criterion is based on the extent to which compelling evidence is provided to support that the proposed project can be completed as planned. Proposals that identify specific potential challenges that may be faced in attempting to conduct the project at each stage of the work and articulate action plans for addressing them should they arise score higher on this criterion.

- *Research / Evaluation environment and team*: Scoring of these criteria are based on the extent to which the PI's research environment can support the proposed project and the evidence provided that the research / evaluation team has a demonstrated record that provides confidence in their ability to do the project. When the target population of the program includes communities of color, the extent to which researchers of color will lead / co-lead the project are considered as well.

    - o *Research environment*
        - Scoring for this criterion is based on the strength of the evidence provided concerning the level of support the PI's institution is able to provide for the proposed project. Considered is the research ranking of the institution in general (similar to the Carnegie Classification of Institutions of Higher Education in which universities are classified by research activity as measured by research expenditures, number of research doctorates awarded, number of research-focused faculty). Resources that would be provided by the institution to support the research / evaluation team while carrying out the proposed project are also considered (e.g., administrative support such as grant support services, office and lab space, personal computers and equipment, software, and /or technological support). Proposals that provide evidence of a high quality research environment and identify specific resources provided by the institution to support the proposed work score higher on this criterion. (Section 5: question 1)

    - o *Expertise of the PI / research team*
        - Scoring for this criterion is based on the strength of the evidence provided (e.g., record of publications, prior grants) that provides confidence in the ability of the PI / Research Team to do the proposed project) Considered when scoring this criterion are evidence of expertise concerning the topic area of the proposed project (e.g., the target population, the intended goals of the program), evidence of expertise concerning the research methods that would be used in the proposed project, and the level to which the intended procedures have been previously pilot tested by the research / evaluation team with the intended target population. Proposals the provide evidence that at least one member of the team (not necessarily the PI) has experience successfully leading projects of similar or greater scope score higher on this criterion (Section 5: questions 2; Section 7: CVs / Resumes of key personnel)

o *Racial/ethnic composition of the leadership of the research team*: BEF is among a growing community of foundations that tracks the diversity of its grantees. In addition, proposed projects that intend to recruit individuals from communities of color to participate in the study must have at least one researcher of color in the leadership (PI / co-PI) level of the research team.

- **ALL** proposals (regardless of the demographics of the target population of the program or the study sample) must provide the race/ethnicity of each of the key personnel of the research team. Each key member of the team must be listed with his/her racial/ethnic identity specified (i.e., do not NOT provide just general descriptions of the diversity of the research team or the institution / organization). (Section 5: question 3a)

- When the leadership of the research team of the proposed project is diverse, the scoring of this criterion is based on the strength of the evidence provided that the collaboration is likely to be successful and that all voices on the team will valued and influence the project. If the leadership team has collaborated in the past, proposals that provide evidence that this collaboration was successful score higher on this criterion. If this is a new collaboration and the team includes a researcher that identifies as white who has collaborated on diverse leadership teams in the past, proposals that provide evidence that this collaboration was successful score higher on this criterion. If this is a new collaboration and the team includes a researcher that identifies as white who has not collaborated on diverse leadership teams in the past, proposals that identify specific ways in which efforts will be made to ensure the collaboration will be successful (e.g., antiracists education opportunities, protocols to address overt and implicit bias) receive higher scores on this criterion. (Section 5: question 3b)

Budget criteria

Scoring for this criterion is based on the extent to which the proposed budget is in line with specific aims / scope of work proposed and is reasonable and justified (Section 1: questions 5-9; Section 6: budget; Section 7: budget justification).

- Considered are the total budget amount requested and whether other support for the project has been secured and/or pending. Also considered is whether all activities proposed are represented in the budget; whether unjustified costs are included; the extent to which the FTE percent requested for each key personnel is reasonable given his/her role and responsibilities (i.e., is neither too high nor too low); and the extent to which estimates for supplies, equipment and other costs (e.g., incentives for participants, costs for assessment materials) are reasonable given the scope of work proposed. Proposals that are able to identify clear links between each cost and the specific tasks needed to complete the proposed work score higher.

- Also considered is whether operational funding for the program to be evaluated has been secured for the period during which the evaluation would occur. All funds requested from BEF must be for evaluation costs (i.e., no program operation costs should be included in the budget). If the proposal includes a subcontract to the organization that operates the program, very clear articulation of the evaluation activities that would be supported by the subcontract should be provided. Proposals that demonstrate that operational funding for the program is already secured (or very likely given past funding) score higher.