UCLA
UCLA Previously Published Works

Title

Gauging treatment impact: The development of exposure variables in a large-scale evaluation study.

Permalink

Journal

ISSN

Authors

Maccalla, Nicole
Purnell, Dawn
McCreath, Heather
et al.

Publication Date

2022

DOI

Peer reviewed

Gauging treatment impact: The development of exposure variables in a large-scale evaluation study

Nicole M. G. Maccalla,

Dawn Purnell,

Heather E. McCreath,

Robert A. Dennis,

Teresa Seeman

University of California Los Angeles, Los Angeles, California, USA

Abstract

While guidance on how to design rigorous evaluation studies abounds, prescriptive guidance on how to include critical process and context measures through the construction of exposure variables is lacking. Capturing nuanced intervention dosage information within a large-scale evaluation is particularly complex. The Building Infrastructure Leading to Diversity (BUILD) initiative is part of the Diversity Program Consortium, which is funded by the National Institutes of Health. It is designed to increase participation in biomedical research careers among individuals from underrepresented groups. This chapter articulates methods employed in defining BUILD student and faculty interventions, tracking nuanced participation in multiple programs and activities, and computing the intensity of exposure. Defining standardized exposure variables (beyond simple treatment group membership) is crucial for equity-focused impact evaluation. Both the process and resulting nuanced dosage variables can inform the design and implementation of large-scale, diversity training program, outcome-focused, evaluation studies.

INTRODUCTION

The Enhance Diversity Study (EDS) is a large-scale, systemic, national longitudinal evaluation of training programs funded by the National Institutes of Health (NIH) Common Fund (McCreath et al., 2017) and managed by the National Institute of General Medical Sciences (NIGMS). Specifically, as outlined by Guerrero et al. in Chapter 1 of this issue, the Diversity Program Consortium (DPC) is made up of three closely integrated initiatives that are working together to achieve the consortium's overarching goals. The Coordination and Evaluation Center (CEC) at the University of California, Los Angeles, was funded to implement the evaluation plan as approved by the DPC Executive Leadership Committee. CEC investigators frame their work as equity-focused impact evaluation (longitudinal, multimethod, quasi-experimental, with a case study component).

Correspondence Nicole M. G. Maccalla, University of California Los Angeles, Los Angeles, CA, USA. nicolemaccalla@gmail.com.

This chapter addresses the intervention operationalization within the evaluation of one of the three initiatives, Building Infrastructure Leading to Diversity (BUILD), which was designed "to implement and study innovative approaches to engaging and retaining students from diverse backgrounds in biomedical research" (NIGMS, 2022, para. 1). The evaluation utilizes national surveys from the Higher Education Research Institute (HERI), CEC annual follow-up surveys, and institutional records to measure outcomes of interest (Davidson et al., 2017). The results from BUILD interventions may yield useful information about what works, for whom, and in what contexts, as well as what doesn't work and why. In the long term, through the dissemination and implementation of effective interventions and strategies, BUILD could have a widespread impact and ultimately enhance the diversity of the NIH-funded research enterprise (NIGMS, 2022).

The unique context and scope of interventions within each BUILD program may highlight training innovations for the biomedical research field. Each of the 10 BUILD awardees is piloting a different approach to enhancing student, faculty, and institutional development (NIGMS, 2022). Understanding the nature of BUILD programming and participation is important for the evaluation study. BUILD programs provide the CEC with program participation data that help define dosage within and across programs (Davidson et al., 2017). These data are crucial because they allow for examination of what specific components of the intervention have null or positive effects and for whom, a central element of an equity-focused impact evaluation (Marra & Forss, 2017).

This critical process information merged with institutional context and individual outcomes data allows for evaluation of the effectiveness of the initiative and consistency of effectiveness of intervention components (Carden, 2017; Derzon, 2018). The extent to which programs can move the needle on intended outcomes sheds light on which intervention components and approaches are effective at producing the outcomes they were designed to produce. Conclusions from the EDS and scalable recommendations for policy and practice rest upon the development of standardized exposure variables, the isolation of effects of various intervention components, and the identification of the factors that influence effectiveness.

The power of experimental approaches in evaluating the effectiveness of educational interventions is well established (What Works Clearinghouse, 2020). The U.S. Department of Education's Institute on Education Sciences designates randomized control trials and quasi-experimental designs as top-tier evidence thresholds in determining what works in education. Central to these designs is the manipulation of treatments based on involvement with an intervention. While there is abundant guidance on how to design rigorous evaluation studies (Alkin & Vo, 2017; Chalmers, 2003; Derzon, 2018; Lewis, Stanick, & Martinez, 2016; Lincoln & Guba, 1986; Venable, Pries-Heje, & Baskerville, 2012) and select well-established outcome measures (Bordage, Burack, Irby, & Stritter, 1998; Jerosch-Herold, 2005; Lee & Pickard, 2013), little is published on how to construct exposure variables that capture nuanced dosage information. Consensus exists that context and process measures must be included in an equity-focused impact evaluation (Carden, 2017; Marra & Forss, 2017), but prescriptive guidance on how to do so is lacking.

This chapter articulates the processes employed in defining the BUILD student and faculty interventions, tracking nuanced participation in multiple programs, and computing the intensity of exposure. Defining standardized exposure variables (beyond simple treatment-group membership) is crucial for conducting an equity-focused impact evaluation, determining the overall effectiveness of the DPC BUILD initiative, and identifying the factors that influence effectiveness.

EVALUATION CONTEXT AND APPROACH

Each BUILD program is designed with a specific context and audience in mind. The theories on which the 10 BUILD programs draw come from evidence-based practices in the field (e.g., culturally responsive mentoring practices, authentic experiences in the classroom, early exposure to research). The DPC model is rooted in what works (and for whom) while also allowing for innovation to materialize to enhance research training environments and potentially support movement toward diversification of behavioral and biomedical health-related sciences. Each BUILD program implements a culturally responsive program; local primary evaluation approaches are informed by use-oriented evaluation theories undergirded by theoretical commitments to social change (see Christie and Wright in Chapter 6). Overall, the CEC's evaluation approach to the EDS can be characterized as context-sensitive (Alkin & Vo, 2017) with influences from various use-oriented (e.g., participatory), methods-oriented (e.g., theory-driven), and value-oriented (e.g., culturally responsive) evaluation frameworks, further detailed in Chapter 1 (Guerrero, et al., this issue) and in Davidson et al. (2017).

DPC and BUILD can be viewed through the larger landscape of diversity, equity, and inclusion (DEI) programming. The focus on tracking clearly defined outcomes for diverse groups in biomedical fields calls for the design and implementation of an equity-focused impact evaluation study that can generate a long-term understanding of changing conditions and differential effects on participants (Marra & Forss, 2017). Sensitive to the challenges of large-scale evaluation and implementation of equity evaluation (Guerrero, et al., this issue), the evaluation team remains particularly mindful of the propensity for averages to hide inequities and that an overemphasis on outcomes and impact could be at the expense of uncovering processes and mechanisms that lead to change (Carden, 2017). The collection of critical process information by way of BUILD program participation data, often seen as unmeasurable in large-scale studies (Carden, 2017), is novel in the field of consortium evaluation.

METHODS AND RESULTS

Given the scope of this national evaluation, the CEC needed to develop a typology of interventions to examine exposure within and across BUILD programs. This typology provides standardized variables for impact analyses and yields a common language and understanding for consortium-wide publications. The process we describe in this section is iterative. Figure 1 provides a snapshot of the three major steps employed in partnership with BUILD programs. Each of these steps is described below, along with the lessons learned. Both the process and resulting nuanced dosage variables can inform the design

and implementation of large-scale, diversity training program, outcome-focused, evaluation studies.

Step 1:   Understand and define the "Program"

Evaluators using a theory-driven approach need to familiarize themselves with the inner workings of the program (i.e., goals, inputs, outputs, activities, outcomes; Chen, 1990). This process enables evaluators to understand and define the program in order to tailor a study's approach and methods to the specific context of the evaluation (Alkin & Vo, 2017). Defining the nature of BUILD student and faculty interventions began in Year 1 and included a comprehensive review of program literature, documents, and artifacts (e.g., program and evaluation proposals, program application materials, websites, reports, presentations) as well as extensive conversations with program personnel. A short description of each program activity was produced in consultation with program staff. Thematic content analysis of these descriptions helped articulate common/standardized activity types (e.g., undergraduate research experience) that apply across BUILD programs. The list of activity types captures both commonality and variability in approaches across programs. Thoroughly understanding a specific activity in any one program and mapping that local activity to a generic or standardized activity type is a labor-intensive process that requires sufficient access to program information, attention to detail, and ongoing collaborative communications with individual BUILD programs.

Preliminary classifications were shared with each of the programs to check for understanding and validity of interpretation. The typology of BUILD interventions and the mapping of individual program activities were compiled in the "BUILD Activity-at-a-Glance,"[1] which serves as a roadmap of how common approaches to student and faculty interventions are uniquely implemented within each site. This program context information is necessary for interpreting values for participation variables in data files. Before each iteration review with programs, the typology and any updates were explained during a consortium-wide presentation with a question-and-answer period. Tables 1 (student only activities), 5.2 (student and faculty activities), and 5.3 (faculty only activities) display the final list of activity types (indicated as Participation Flags) and descriptions used in the EDS. It should be noted that not all BUILD programs offer each activity type. For example, each BUILD program offers the "Scholar" opportunity, but only six offer the "Associate" opportunity. For a detailed description of what participation in BUILD programs looks like, please consult the forthcoming article by Maccalla and Purnell.

Over time, the typology of BUILD interventions was refined based on: (a) how the interventions were actually delivered (e.g., Diversity Training began as an activity type but was pulled into the Learning Community activity type because that is where the activity occurred and that is what is trackable); (b) programmatic changes (e.g., Phase I and Phase II of the programs, COVID-19 pandemic impacts); and (c) availability of individual-level participation data. Highly specialized program activities such as tutoring for a specific course or activities that prove infeasible to track well, such as peer mentoring, are noted

---

[1]The "BUILD Activity-at-a-Glance" can be found at: https://www.diversityprogramconsortium.org/flles/view/docs/Consortium-wide_Data_Documentation_DPC_BUILD_Activity-at-a-Glance.xlsx

in narrative descriptions of the programs but excluded from mapping to activity types. The typology captures participant designations (e.g., Mentor), types of funding support (e.g., Mentor Financial–Materials and Supplies), and engagement in specific program activities (e.g., Mentoring of Undergraduate Students). Multiple member-checking strategies at various stages ensured the accuracy of mapping and completeness of activity lists. Thinking about individual programs in these universal terms can be a challenge. At times, it can be difficult for program representatives to take the meta-view of representing their unique intervention components in the typology of BUILD interventions.

This significant scope of work is possible because of the well-resourced evaluation budget, working relationships with BUILD partners, and the technical capabilities of the CEC. The mapping process is ongoing as programs themselves continue to evolve. A clear typology of BUILD interventions is vital but so are ongoing communications with programs to make sure the evaluation team is "getting it right." Validity of classification is imperative, not only in being able to produce accurate descriptions of programming (and variability of participation) but also in informing the collection of data elements associated with each activity (e.g., number of hours associated with the training). If individual program activities are incorrectly mapped, important data elements may be missed (and may be unretrievable), which can affect the quality of the impact analyses. The interdependencies between the qualitative and quantitative aspects of this work cannot be overstated.

Step 2: Develop and implement tools to track participation in program activities

Participation information reported directly by the program is critical process information for the evaluation. It allows for a range of effectiveness to be discerned, which contributes to understanding how outcomes are achieved. The CEC developed the DPC Tracker, a relational database-backed web application suite engineered around requirements for a secure, centrally hosted system that can support three different classes of users in the cooperative agreement: the funder (NIH/NIGMS), the external evaluator (the CEC), and each of the 10 consortium member sites (BUILD programs). Figure 2 depicts the core entities in the Tracker. For an individual BUILD program, it provides a mechanism to declare the scope of activities that make up the program. Program activities (e.g., Proposal Writing Workshop) are defined. Each offering is represented as a session (e.g., Proposal Writing Workshop, Spring 2021).

Attendance rosters are the mechanism by which the Tracker associates participants with sessions. These rosters capture the point-in-time attributes, names, and emails of individual participants. The specialized data elements collected by the CEC for each Participation Flag are listed in Tables 1 (student only activities), 5.2 (student and faculty activities), and 5.3 (faculty only activities). Rich details related to undergraduate research experiences, conference attendance/participation, novel curriculum, length of training, and funding levels are collected to account for variability in experiences across individuals, even within the same program.

This tracking system requires maintenance and refinement, including a staff member directly dedicated to processing participation data. BUILD programs are encouraged to upload session-level rosters as close to the date of the activity as possible. Most programs have

found that uploading rosters in waves—often connected to major program milestones, such as the end of the semester—is most efficient. The CEC processes the participation data at regular intervals and releases session lists mapped to Participation Flags back to programs for member checking on an annual to semi-annual basis. These status reports assist with quality control.

Unique individual identifiers in the system are names and emails. Email is included to facilitate data collection with program participants and for the association to outcomes data. Because both students and faculty may use multiple email addresses, multiple records can be generated for the same individual. Thus, the Tracker includes a function allowing records for the same individual with multiple email addresses to be combined. This ensures that all program participation for a given individual is tracked. Critical to the work of classifying participation in program activities is a report in which each observation represents an individual participant in an activity session for later compilation.

A multistage roll-out process allowed for the incorporation of stakeholder feedback into the application and to ensure use. Even with stakeholder buy-in, the general conceptual structure of how program interventions are represented in the Tracker can be a challenge for individual BUILD programs. The typology structure of individual participants within sessions within activities, as well as the relationships of activities to Hallmarks of Success (outcomes) and other constructs, takes time to appreciate and navigate. It has taken valuable time and energy to communicate, train, and assist partners in understanding this aspect of the national evaluation.

When BUILD program staff are onboarded to work with the Tracker, they create a Tracker user account, attend Tracker training sessions, and access Tracker interface resources (i.e., "how-to videos" for session creation and roster upload instructions). They have access to consortium-wide presentations on the purpose and use of the Tracker and on the development of exposure variables in the EDS. They also receive an overview of naming conventions, an orientation to standardized activity types (Participation Flags), and information on specialized data elements that have been requested, with corresponding Tracker templates.

The investment of time and resources into the development and maintenance of the Tracker and the collection of participation data is worthwhile. There are multiple threats to internal validity that can be addressed with such process data. First, by knowing who participated in the intervention, we are able to examine differences due to systematic bias in the selection of program participants. Related to this is documentation of attrition, in the event participants are not exposed to all planned program components because they have stopped participating. It also assists in understanding generalized approaches to the interventions by allowing examination of different implementation strategies across programs. Finally, program reports of participants promote understanding of the temporal ordering of exposure and achievement of outcomes.

Because these data will be used in evaluation analyses, monitoring data quality is important. Data quality can be impacted by a lack of understanding of program activities and features,

unavailability of session rosters, incomplete information in session roster uploads, changes in personally identifiable information shared, time and resource constraints, staff turnover, a failure to share participation data with the CEC in a timely fashion, a failure by the CEC to process data in a timely fashion, and a lack of program or evaluator responsiveness. Significant time is spent adjudicating data inconsistencies to ensure the validity of raw participation data. To some extent, data collection and processing must take place over a sustained period before the evaluation team realistically has a good handle on what is possible regarding the development of exposure variables. The successful give and take between individual BUILD programs and the CEC hinges upon trust, dedication, collaboration, and good working relationships.

Step 3:   Create exposure variables and gauge dosage

By the end of Year Six, a working version of the typology of BUILD interventions was in place for the EDS, and program activities and sessions to date had been mapped. The focus shifted to using information uploaded to the Tracker to develop measures of exposure to BUILD program activities for use in impact evaluation analyses: What is the effect of x on y? What works, for whom, and under what conditions? The phased rollout of increasingly sophisticated versions of exposure variables is well underway. This section provides an overview of what is already available to analysts and what is still planned for release.

To support analyses, we created measures of participation by using the export report from the Tracker and applying the activity mapping to that raw data. That allowed us to create a series of indicators of whether or not each individual in the EDS participated in an activity of that type (see Tables 1, 2, and 3 for standardized activity types). Participation data are currently summarized at the person level with three variables for each standardized activity type: (a) a Participation Flag, which is a dichotomous indicator of involvement (ever participated in an intervention in that category–No/Yes); (b) the start date for the activity (date of the earliest involvement in an intervention in that category); and (c) the end date for the activity (date of the latest involvement in an intervention in that category). Every 3–6 months, the CEC releases an updated version of these data, including the unique ID for each of the student and faculty participants, enrolled in the study. This can be merged with any student and/or faculty survey or outcomes data set to support impact analyses.

The Participation Flag variables are the most basic version of exposure variables, indicating if a student or faculty member *ever participated* in *any activity* that maps to the Participation Flag. For example, a value of "1" for PEDGTRAIN (the flag for Pedagogical Training) indicates that the faculty member has attended at least one BUILD-sponsored Pedagogical Training session at some point since the beginning of the grant. Inspection of start and end dates associated with the PEDGTRAIN variable helps the analyst estimate timing of the intervention to best select pre–post-survey data sets and constructs of interest (e.g., instructor self-efficacy). This structure enables time-dependent modeling of involvement when examining outcomes. The main limitation of the binary format of exposure variables is that it masks the extent of exposure. The fact that a faculty member who attended a single, 1-h pedagogical training session has the same value for the variable as a faculty member who regularly attended monthly pedagogy workshops over a several-year period illustrates

this point. The primary strength of the simplest version of exposure variables is that usage is intuitively clear.

More sophisticated releases of exposure variables will include frequency of involvement, total hours of involvement, and nature of involvement in specific activity types. Additionally, a composite variable capturing the extent of involvement (high, medium, low) in a collection of student or faculty training activities will be computed. For all the Participation Flags listed in Tables 1, 2, and 3 that represent participation in recurring program activities, frequency of participation—a total count of attendance at unique activity sessions—will be made available to analysts. For any Participation Flag representing training to enhance research skills, teaching, or mentoring practices, total hours involved in training sessions will be made available to analysts.

Variables indicating the frequency of involvement and total hours of training will help address the question of the extent of exposure to BUILD interventions. Examination of start and end dates for these variables will assist analysts in correctly modeling the timing of various interventions and determining activity, program, and initiative impact. For program activities that unfold in a variety of ways (e.g., participation in scientific conferences), even with similar frequency of involvement, differential effects on participants may be observed. For variables such as CONF (i.e., whether a student was a presenter or attendee at a local or national conference), the nature of involvement will be important to consider when isolating the effects of intervention components on diverse participants.

Each BUILD program offers different types of activities and packages the student and faculty experience uniquely; some activities are compulsory and others are optional. Being able to look across BUILD programs and construct comparable treatment groups (i.e., high/medium/low exposure) helps determine program and initiative effects on individuals in longitudinal analyses; it may even aid in the identification of leverage points or minimum thresholds of participation that produce desired results. If differences between programs are uncovered, analysts can then focus on explaining those differences based on program theory and innovative implementation features. Uncovering particularly impactful approaches to BUILD programming will support the generation of quality recommendations for colleges and universities implementing student and faculty training programs.

It is estimated that by the end of Year Eight (in mid-2022), nuanced exposure and dosage variables will be available for a more detailed understanding of BUILD program interventions. A technical report detailing the computation of nuanced exposure variables will be published on the DPC website and should complement this description of the conceptual approach and process of developing exposure variables vital for equity-focused impact analyses.

CONCLUSION AND IMPLICATIONS

Even though the body of literature determining the impact of various educational interventions is growing, the process of developing exposure variables in large-scale evaluation studies is rarely discussed. A better understanding of the efforts involved in

accurately reflecting participation in program activities can support evaluation planning and management processes. Grasping variations in program activities enables evaluators to test how variations in delivery and participation affect intended outcomes (Derzon, 2018), which may uncover what factors influence effectiveness.

Developing methods for standardizing the measurement of program participation while maintaining variability in exposure supports successful equity-focused impact evaluation through the attention to process and the preservation of program-level and individual-level context information. Merging nuanced participation data with measures of institutional context and valid and reliable survey constructs can support the examination of questions such as Does it work? How and in what ways does it work? For whom does it work? What doesn't work, and why? Consistent definitions of BUILD exposure and the development of well-calibrated exposure variables ensure reliable precision in consortium-wide evaluation results. Findings from the EDS will ultimately inform national policy in higher education and the successful transference of effective, scalable, evidence-based practices to diversify behavioral and biomedical health-related sciences.

Drawing on core principles, such as partnership and collaboration, while paying attention to context, process, and mechanisms for change when evaluating program impacts can serve as a way to "infuse principles of equity evaluation into more traditional evaluation studies" (Guerrero, et al., this issue, p.). Even if the implementation of evaluation theories focused on diversity, equity, and inclusion seems out of reach, equity-focused impact evaluation should not be. The process of developing exposure variables in the EDS, as described in this chapter, offers a roadmap for unpacking process and context measures in an outcome-focused evaluation study of a large-scale training program.

Acknowledgments

Biographies

Nicole Maccalla, PhD, is housed in UCLA's School of Education and Information Studies and serves as a lead investigator for the Coordination and Evaluation Center's Enhance Diversity Study, as well as the BUILD evaluation study coordinator and chair of the Evaluation Implementation Working Group.

Dawn Purnell, MA, is housed in the David Geffen School of Medicine at UCLA and serves as a staff researcher for the Coordination and Evaluation Center's Enhance Diversity Study and co-chair of the Program Implementation Working Group.

Heather McCreath, PhD, is an adjunct professor in the Division of Geriatrics in the David Geffen School of Medicine at UCLA and serves as co-director of the Data Coordination Core in the Coordination and Evaluation Center for the Diversity Program Consortium.

Robert A. Dennis, PhD, is the director of software development for the David Geffen School of Medicine at UCLA and the Computing Technologies Research Lab and he is a member of the Data Coordination Core responsible for informatics systems supporting the Coordination and Evaluation Center for the Diversity Program Consortium.

Dr. Teresa Seeman, PhD, is a professor of medicine and epidemiology at the David Geffen School of Medicine at UCLA and serves as the co-principal investigator for the Coordination and Evaluation Center and co-director of the Data Coordination Core for the Diversity Program Consortium.

REFERENCES

Alkin MC, & Vo AT (2017). Evaluation essentials: From A to Z (2nd ed.). Guilford Press.

Bordage G, Burack JH, Irby DM, & Stritter FT (1998). Education in ambulatory settings: Developing valid measures of educational outcomes, and other research priorities. Academic Medicine, 73(7), 743–750. 10.1097/00001888-199807000-00009 [PubMed: 9679462]

Carden F (2017). Building evaluation capacity to address problems of equity. New Directions for Evaluation, 154, 115–125. 10.1002/ev.20245

Chalmers I (2003). Trying to do more good than harm in policy and practice: The role of rigorous, transparent, up-to-date evaluations. The ANNALS of the American Academy of Political and Social Science, 589(1), 22–40. 10.1177/0002716203254762

Chen HT (1990). Theory-driven evaluations. Sage Publications.

Davidson PL, Maccalla NMG, Abdelmonem AA, Guerrero L, Nakazono TT, Zhong S, & Wallace SP (2017). A participatory approach to evaluating a national training and institutional change initiative: The BUILD longitudinal evaluation. BMC Proceedings, 11(Suppl. 12), 157–169. 10.1186/s12919-017-0082-9

Derzon JH (2018). Challenges, opportunities, and methods for large-scale evaluations. Evaluation & the Health Professions, 41(2), 321–345. 10.1177/0163278718762292 [PubMed: 29577743]

Jerosch-Herold C (2005). An evidence-based approach to choosing outcome measures: A checklist for the critical appraisal of validity, reliability and responsiveness studies. British Journal of Occupational Therapy, 68(8), 347–353. 10.1177/030802260506800803

Lee TA, & Pickard AS (2013). Exposure definition and measurement. In Velentgas P, Dreyer NA, Nourjah P, Smith SR, & Torchia MM (Eds.), Developing a protocol for observational comparative effectiveness research: A user's guide (pp. 45–58). Agency for Healthcare Research and Quality. Retrieved from: https://www.ncbi.nlm.nih.gov/books/NBK126191/

Lewis CC, Stanick CF, & Martinez RG (2016). The Society for Implementation Research Collaboration Instrument Review Project: A methodology to promote rigorous evaluation. Implementation Science, 10, 1–18. 10.1186/s13012-014-0193-x

Lincoln YS, & Guba EG (1986). But is it rigorous? Trustworthiness and authenticity in naturalistic evaluation. New Directions for Program Evaluation, 30, 73–84. 10.1002/ev.1427

Marra M, &Forss K (2017). Thinking about equity: From philosophy to social science. In Forss M (Ed.), Speaking justice to power (pp. 21–33). Routledge.

McCreath HE, Norris KC, Calderón NE, Purnell DL, Maccalla N, & Seeman TE (2017). Evaluating efforts to diversify the biomedical workforce: The role and function of the Coordination and Evaluation Center of the Diversity Program Consortium. BMC Proceedings, 11(Suppl. 12), 15–26. 10.1186/s12919-017-0087-4 [PubMed: 29375658]

National Institute of General Medical Sciences. (2022). Building Infrastructure Leading to Diversity (BUILD) Initiative. Retrieved from: https://www.nigms.nih.gov/training/dpc/pages/build.aspx

Venable J, Pries-Heje J, & Baskerville R (2012). A comprehensive framework for evaluation in design science research. InPeffers K, Rothenberger M, &Kuechler B (Eds.), DESRIST: Design science research in information systems: Advances in theory and practice (pp. 423–438). Springer.

What Works Clearinghouse. (2020). What Works Clearinghouse standards handbook, Version 4.1. U.S. department of education, institute of education sciences, national center for education evaluation and regional assistance. Retrieved from: https://ies.ed.gov/ncee/wwc/handbooks
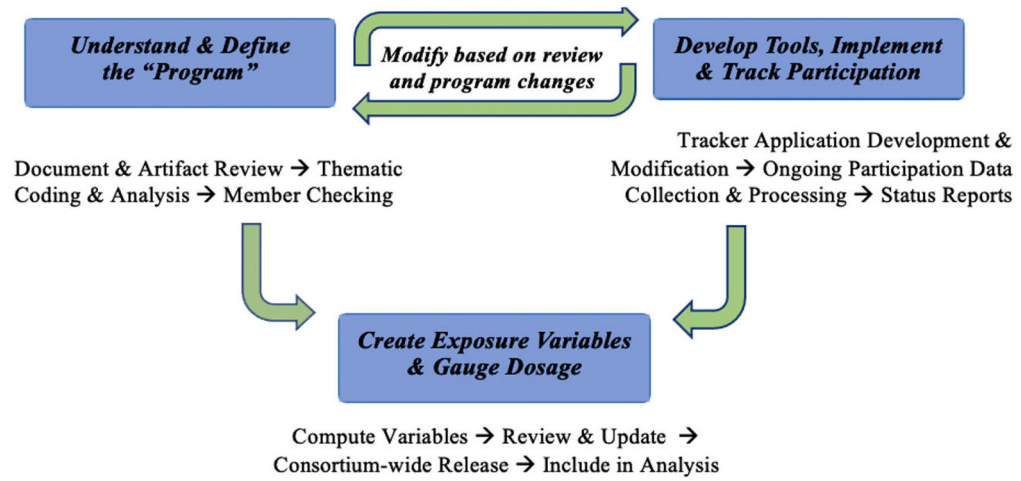
FIGURE 1.
Overview of iterative 3-stage process to develop exposure variables in the enhance diversity study
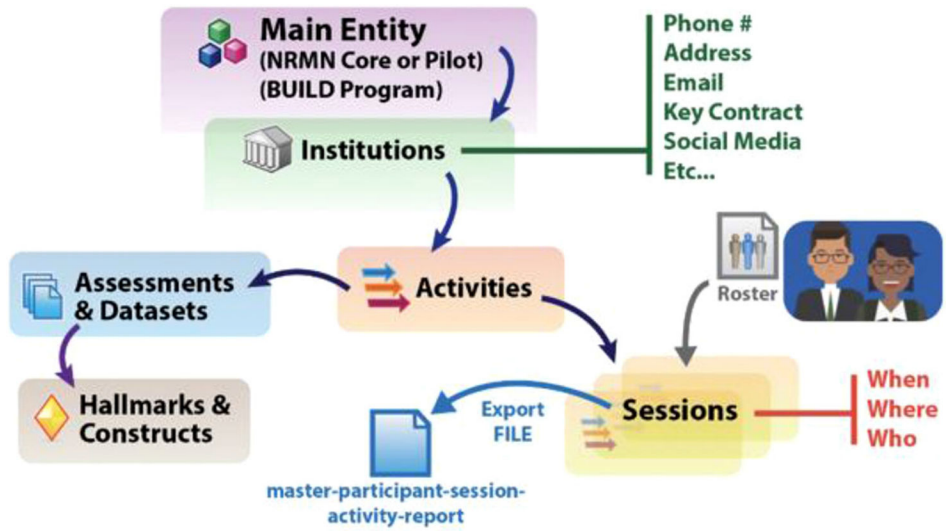
FIGURE 2.
Core data entities in the enhance diversity study tracker

TABLE 1

Participation flags that map to BUILD student activities, with descriptions and identification of specialized data elements

Participation flag	Description	Specialized data elements
Scholar designation	Program-defined, most intensely treated and supported group of students. Scholars often receive tuition support or stipend, research training, and mentorship. Compulsory and structured participation in a host of BUILD activities is common for this group.	None
Associate designation	Program defined, less intensely treated and supported group of students, often participating in a subset of structured BUILD activities. Some programs recruit *Scholars* from the *Associate* pool.	None
TL4 funded	NIH funding source. Often includes students who receive tuition scholarship.	None
RL5 funded	NIH funding source. Often includes students who receive stipends or salaries.	None
Novel curriculum–Enrollment	Includes enrollment/attendance in a new BUILD-sponsored course or existing course revamped for BUILD purposes (likely maps to NCDEV list for faculty).	Course name, course number, units
Career advancement and development	Includes networking opportunities, GRE prep, fieldtrips, career panels, career speaker series, underrepresented group professional exposure, grad school application assistance, career advising, CV development workshop, interviewing skills, test prep, etc.	None
Learning community participant	Participation in a mandatory BUILD Learning Community, usually on a weekly/biweekly/monthly meeting schedule. BUILD students are often required to attend, for zero credit or low credit. Often includes diversity training (e.g., bias, stereotype threat).	None
Undergraduate research experience	Includes BUILD-affiliated student-directed research and/or mentored undergraduate research experiences during the academic and/or summer term(s).	Lab Location (on campus prime, on campus partner, off campus lab), Length of Lab Placement (summer, semester, academic year), Total Hours
Summer bridge	Participation in a summer program focused on academic preparation (e.g., writing, statistics) and orientation to the university, typically for entering freshman and/or transfer students.	None

Note. Standard data elements collected from every activity session roster upload include: full roster of participants, including first and last name, ID, ID type, and email; role (typically a faculty or student designation); and start and end date of the activity session.

TABLE 2

Participation flags that map to BUILD student and faculty activities, with descriptions and identification of specialized data elements

Participation flag	Description	Specialized data elements
Research training and support	Includes workshops or trainings focused on developing research skills (e.g., CITI, ethics, lab safety, working with specialized equipment, analyses, proposal development, publication of findings) or devoted to research topics (e.g., invited research speaker series).	Hours
Conference participation	Includes BUILD-supported conference attendance and/or presentation at BUILD, local, and national conferences.	Name of conference, Scale (BUILD or campus sponsored, regional/national/international), Role (attendee, presenter)
Other Funding Support	Received other funding support from the BUILD program not already designated (e.g., travel award, stipend, publication costs).	None

Note. Standard data elements collected from every activity session roster upload include: full roster of participants, including first and last name, ID, ID type, and email; role (typically a faculty or student designation); and start and end date of the activity session.

TABLE 3

Participation flags that map to BUILD faculty activities, with descriptions and identification of specialized data elements

Participation flag	Description	Specialized data elements
Pedagogical training	Includes participation in pedagogical training aimed to enhance the quality of teaching and learning (particularly for underrepresented students). Format may include pedagogy workshops, pedagogy training, pedagogy modules, pedagogy best practices, culturally responsive pedagogy models, pedagogy panels, pedagogy colloquiums and symposiums, etc.	Hours
Novel curriculum–development	Includes the active development of a new or existing course with BUILD support. Courses are typically developed once and taught/offered many times (likely maps to NOVELCURR list for students). Highest priority for NCDEV is the developer(s) of the course(s).	Course name, course number, units, course description
Mentor training	Includes participation in mentoring workshops, mentor training, mentor modules, mentoring best practices, culturally responsive mentorship models, mentor panels, mentors colloquiums, National Research Mentoring Network (NRMN) mentor training, etc.	Hours
Mentor (to undergraduate, graduate, postdoc)	Serving as a mentor for BUILD students – includes academic, career, and/or research mentoring of students (including mentored undergraduate research experiences) and/or junior faculty members.	Number of Mentees, BUILD Mentee Names, Mentoring Relationship (Academic, Career, Research), Mentee Level (UG, Postdoc, Grad)
Mentor financial support	Received funding to support mentoring, often looks like mentor materials and supplies.	None
Pilot project award ($10K–$50K)	Received BUILD-sponsored Pilot Project funding (most likely through a competitive application/review process), often referred to as "seed money" for research.	Amount, abstract
Lab grant award	Received Lab Grant for small instruments or equipment.	Amount, equipment
Release time and/or summer support	Received salary support for release time or summer support.	None

Note. Standard data elements collected from every activity session roster upload include: full roster of participants, including first and last name, ID, ID type, and email; role (typically a faculty or student designation); and start and end date of the activity session.