PennState

The Methodology Center

ADVANCING METHODS, IMPROVING HEALTH

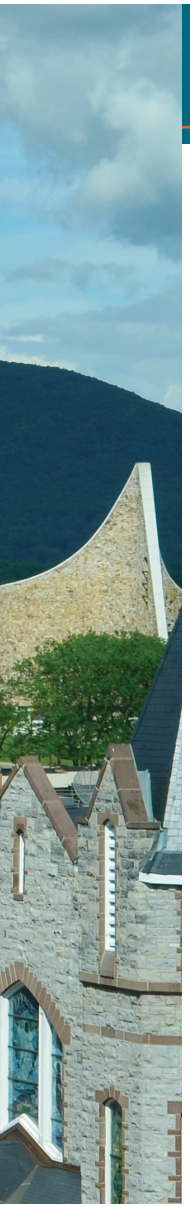
Annual Report

2015-16

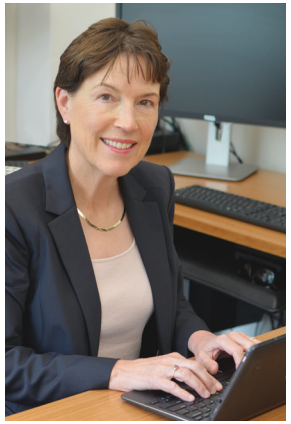
Our mission

is to advance public health by improving experimental design and data analysis in the social, behavioral, and health sciences.

We draw upon and integrate methodological perspectives from a variety of disciplines, including statistics, engineering, psychology, and human development, to develop new quantitative methods for research focusing on vital public health issues, especially drug abuse and HIV.



From the Director



Welcome! It has been an exciting year: we have a new center grant, a new physical space, and new research projects. All of this change, however, is in service to the same mission. We are excited about where these new opportunities will take our work in the coming year.

We recently were awarded a new National Institute on Drug Abuse P50 Center of Excellence grant, "Complex Data to Knowledge." This new grant is helping us focus our research on analytic and experimental design methods for complex data. Technological advances are allowing researchers to collect vast amounts of data in many new formats. New methods are needed to extract the greatest amount of knowledge from the data, and developing these methods is the focus of our new grant. Read more on page 8.

Our Penn State offices also have a new physical home. We now occupy the fourth floor of the beautiful new Health and Human Development Building on Penn State's University Park campus.

The image to the left is the view from our new high-tech conference room. See the back cover for more.

Most exciting of all, of course, are our research projects. The new center grant and several other new grants are helping us leverage complex data to address issues related to gambling addiction, risky sex and drinking, mHealth, HIV prevention, and drug abuse.

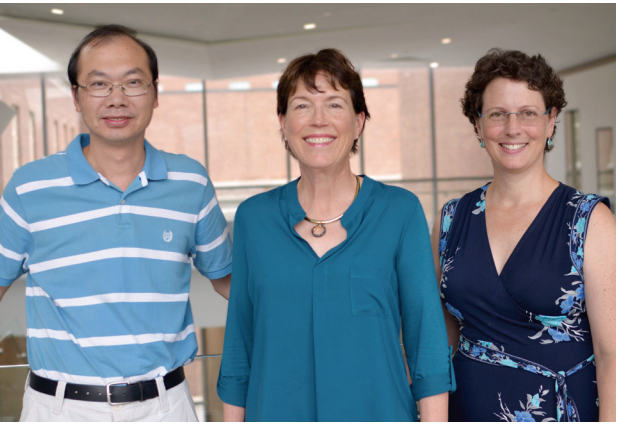
I hope you'll take a few minutes to learn a little more about our research.

Enjoy!

Linda M. Collins
Director, The Methodology Center
Distinguished Professor, Human Development and Family Studies
Professor, Statistics

TO LEARN MORE ABOUT THE METHODOLOGY CENTER, VISIT **methodology.psu.edu**

Twenty Years of LCA



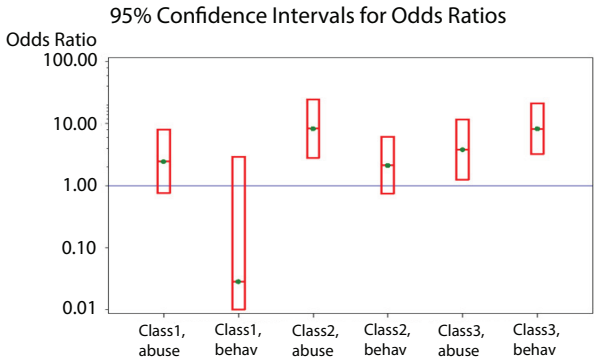
Runze Li, Linda Collins, and Stephanie Lanza

Since the inception of The Methodology Center, we have researched latent class analysis (LCA). In the coming years, we will explore using LCA to model complex moderators, and we will develop new models for latent class variables that can change across time. As the LCA project evolves, we reflect on two decades of LCA research at The Methodology Center.

What is LCA?

LCA allows researchers to discover unobservable subgroups within a population. A variable that cannot be directly measured (i.e., a latent variable) can help us understand a population better. For example, we cannot directly measure a young person's risk, but the combination of measurable variables—such as household income, neighborhood poverty, and history of abuse—can be used to create latent classes that describe different risk profiles within the population. Understanding those profiles, in turn, can help researchers understand where to focus intervention efforts (Lanza, Rhoades, Nix, Greenberg, & CPPRG, 2010).

Figure created using Methodology Center LCA graphics software.



Methodological Contributions

LCA has existed for decades, but Linda Collins was one of the first researchers to adopt and expand LCA for use in the behavioral, health, and social sciences. Early Methodology Center research on latent class models led to the development and application of latent transition analysis (LTA), which describes changes in latent class composition over time (Velicer, Martin, & Collins, 1996). LTA allowed for many new insights and provided a framework for testing the gateway drug hypothesis (Collins, 2002). Later Methodology Center research led to other extensions of LCA including a novel approach to LCA with a distal outcome (Lanza & Rhoades, 2013), which allows researchers to describe associations between latent class membership at one time point and outcomes later in life. Some of our recent work has integrated LCA with causal inference. This work extends LCA so that researchers can understand the causal impact of a predictor on latent class membership (Butera, Lanza, & Coffman, 2014). In order to make LCA and LTA more accessible to drug abuse and HIV researchers, Linda Collins and Stephanie Lanza wrote the book on these methods, *Latent Class and Latent Transition Analysis: With Applications in the Social, Behavioral, and Health Sciences* (2010).

Building Software Tools

Methodology Center scientists have developed several free software packages to enable researchers to estimate latent class models. In 1996, Methodology Center researchers created WinLTA, a standalone Windows program for estimating LCA and LTA models. In 2008, we released the more powerful SAS procedures PROC LCA and PROC LTA. These procedures have been downloaded thousands of times and continue to be downloaded and used daily by researchers around the world. We frequently expand and update the capabilities of the procedures so that researchers can address increasingly complex research questions. In parallel, we developed six SAS macros to expand PROC LCA's functionality. The macros allow users to generate graphics displaying their results, to perform the bootstrap likelihood ratio test for LCA, and to estimate the association between latent class membership and a distal outcome. In 2013, we developed the LCA Stata Plugin to enable Stata users to apply these techniques to their data.

Training Scientists

Methodology Center scientists have consistently trained researchers to apply LCA and related methods. Our scientists have offered dozens of intensive workshops around the world at universities, private companies, and conferences. We maintain a wide array of educational resources on our website, including podcasts, FAQs, reading lists, and explanations of LCA applications and the LCA mathematical model.

The Future of LCA

As The Methodology Center transitions into our new funding period, our work with LCA is evolving. We are embarking on a project to integrate LCA with time-varying effect modeling. (Read more on page 8.) At the same time, LCA is being applied far beyond The Methodology Center. LCA has led to new ways to understand data and to target interventions, and scientists around the world continue to apply latent class models to new data sets. LCA's impact on public health will continue to unfold for many more years.

REFERENCES

Butera, N.M., Lanza, S.T., & Coffman, D.L. (2014). A framework for estimating causal effects in latent class analysis: Is there a causal link between early sex and subsequent profiles of delinquency. *Prevention Science, 15*(3), 397-407.

Collins, L.M. (2002). Using latent transition analysis to examine the gateway hypothesis. In D. Kandel & M. Chase (Eds.), *Examining the gateway hypothesis: Stages and pathways of drug involvement* (pp. 254-269). Cambridge, UK: Cambridge University Press.

Collins, L.M., & Lanza, S.T. (2010). *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Hoboken, NJ: John Wiley & Sons.

Lanza, S.T., & Rhoades, B.L. (2013). Latent class analysis: An alternative perspective on subgroup analysis in prevention and treatment. *Prevention Science, 14*(2), 157-168.

Lanza, S.T., Rhoades, B.L., Nix, R.L., Greenberg, M.T., & CPPRG (2010). Modeling the interplay of multilevel risk factors for future academic and behavior problems: A person-centered approach. *Development and Psychopathology, 22*(2), 313-335.

Velicer, W.F., Martin, R.A., & Collins, L.M. (1996). Latent transition analysis for longitudinal data. *Addiction, 91*(12s1), 197-210.

Productivity and Achievements

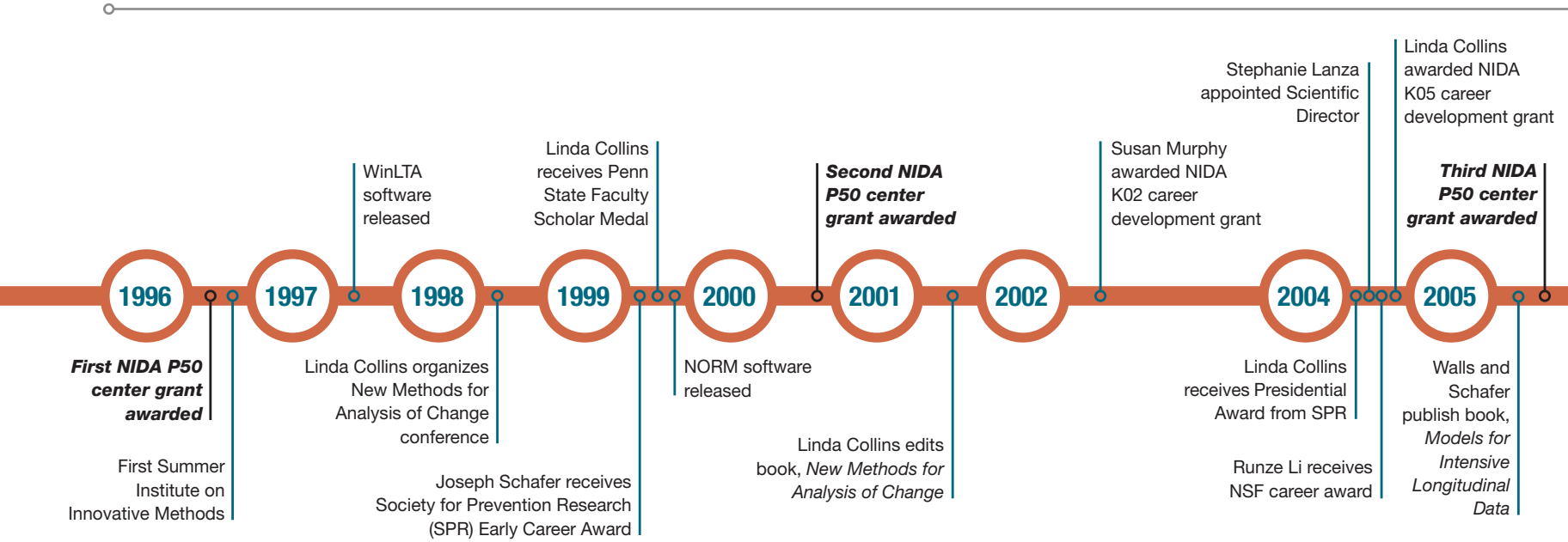
10 Average number of downloads per day of Methodology Center software since 2011

14,000 Free PROC LCA and Stata LCA downloads since 2011

950 Conference presentations since 1996

65 Book chapters published since 1996

225 Peer-reviewed articles published in the last 5 years



Our top 5 articles have been cited more than

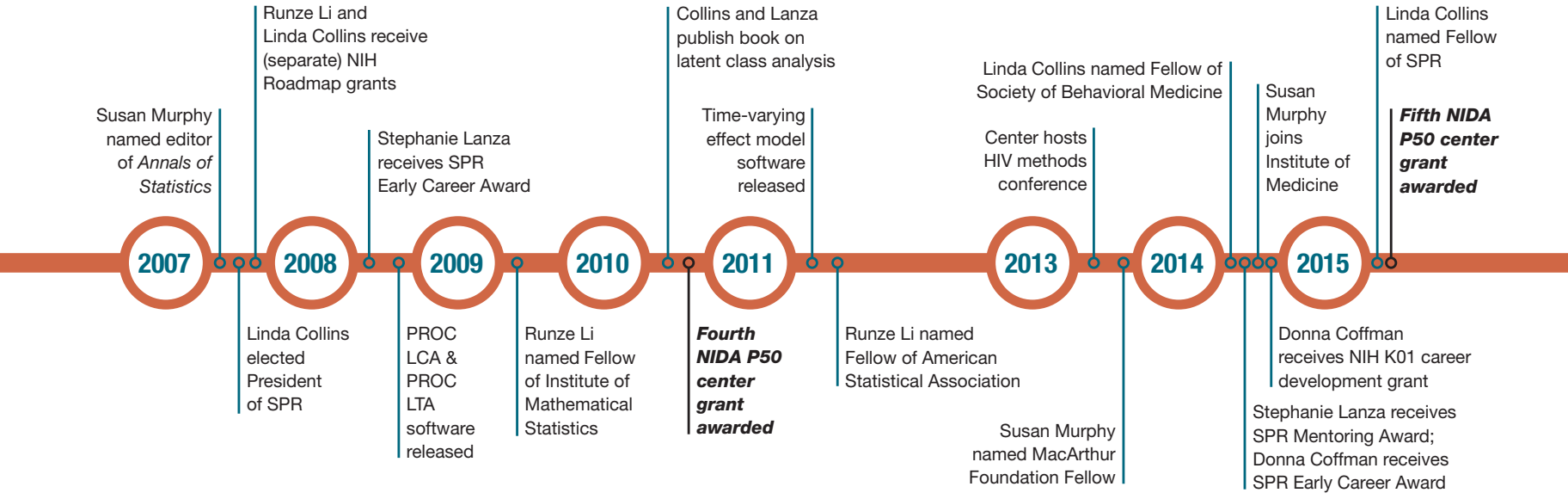
10,000×

33,773
web visitors in 2011

4
Open NIH program announcements that specify an interest in MOST- or SMART-related designs

137,768
web visitors in 2014

25+
Studies using SMART to build adaptive interventions



Advancing Methods for Complex Questions

Smartphones and biosensors collect data at a rate and volume that could scarcely be imagined a few decades ago. Genetic data can contain more than 10,000 variables per subject. These and other emerging technologies are producing new, rich data sets, and many new studies are collecting data from multiple sources that must be integrated. These complex new data can inform a new generation of interventions for drug abuse and HIV prevention. However, statistical analysis methods, the keys researchers use to unlock the scientific knowledge contained in behavioral data, have not kept up with the complexity of modern data sets or the sophistication of the questions posed by behavioral researchers. Several new grants at The Methodology Center seek to capitalize on the information contained in these complex data to allow researchers to create more effective and efficient behavioral interventions.

Susan Murphy, courtesy of the John D. and Catherine T. MacArthur Foundation

Using funding from these new grants, Methodology Center researchers will develop and disseminate innovative statistical methods and apply these new methods to pressing questions in drug abuse and HIV research. The cornerstone of Methodology Center funding is a new Center of Excellence grant from the National Institute on Drug Abuse (NIDA).

Intervention in Your Pocket: The Just-in-Time Adaptive Intervention (JITAI)

NIDA Center of Excellence Grant – Research Project 1

The ubiquity of mobile devices like smartphones has created new possibilities for helping people curb a broad spectrum of unhealthy behaviors. This research project will develop learning algorithms, statistical methods, and clinical trial designs to make the best use of mobile technologies to deliver personalized behavioral interventions. These interventions are known as just-in-time adaptive interventions (JITAIs, pronounced like "Jedis" from *Star Wars*). JITAIs link real-time data about an individual, like their social context, feelings, or behavior, to the best treatment option, like a behavioral intervention, a cognitive intervention, or social support. JITAIs, a form of precision medicine, respond to each individual's specific circumstances.

This research will lead to the construction of effective, individualized, mobile interventions. JITAIs integrate ideas from statistics, computer science, and behavioral science to enable scientists to construct more effective interventions to prevent drug abuse and HIV.

MacArthur Fellow Susan Murphy leads this project. Susan and her team are developing the decision rules needed to construct JITAIs. To do this, they are expanding on the concepts they developed for the sequential, multiple assignment, randomized trial (SMART). SMART is being used in dozens of projects around the world to build adaptive interventions for health problems ranging from autism to drug abuse to obesity. Methodology Center research on SMART continues under the guidance of Susan's colleagues and former trainees, Daniel Almirall and Billie Nahum-Shani.

Integrating Methods to Understand how Interventions Work: LCA and TVEM

NIDA Center of Excellence Grant – Research Project 2

Stephanie Lanza will lead research to integrate methods from two existing Methodology Center research projects: latent class analysis (LCA) and time-varying effect modeling (TVEM). LCA allows researchers to detect unobserved subgroups within a population. TVEM allows researchers to model complex relationships without assuming the relationships are constant over time. Combining these approaches will allow researchers to investigate previously unanswerable questions, such as "Does gender differentially moderate the effect of drug abuse services for people of different ages?" and "During a smoking quit attempt, *when* does baseline dependence moderate the effect of a smoking cessation intervention on withdrawal symptoms?" Answering these types of questions will help researchers understand the way that risk factors and intervention effects vary over time for subgroups of the population.

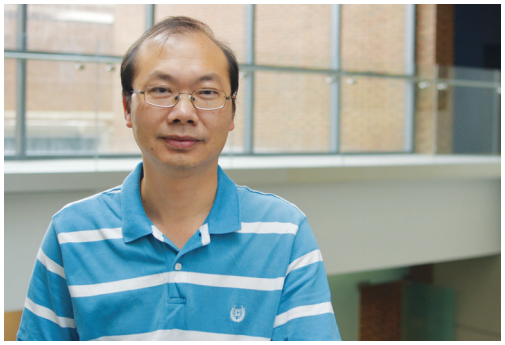
The ultimate goal of this research is to develop new tools that allow HIV prevention and drug abuse researchers to discover clinically useful moderators with complex relationships to intervention effects. This will inform the development of more effective, targeted interventions. The researchers will apply these new techniques to investigate the complex relationships within a large study of addiction services, a randomized smoking cessation intervention, and a youth risk intervention.

Richer Data, Greater Challenges: Variable Selection for High Dimensional Data

NIDA Center of Excellence Grant – Research Project 3

Drug abuse and HIV researchers are increasingly collecting ecological momentary assessments (EMA) and genetic data. However, in EMA, genetic data, and integrations of the two, there may be hundreds of subjects and hundreds of thousands of variables. Paradoxically, this abundance of variables has a crippling effect on exploratory data analyses because nearly all available analytic procedures break down when the number of variables exceeds the sample size.

This project will improve methods for variable screening, which will allow researchers to reduce the number of predictors to a subset that have a potential impact on the outcome of interest. For several years, Runze Li and his collaborators have been developing and extending methods



Runze Li

for variable screening and selection. Runze has also worked on developing methods, such as time-varying effect modeling (TVEM), for the analysis of intensive longitudinal data. This new line of research will incorporate TVEM and variable screening so that researchers can answer questions such as "Which genetic, individual, and social factors predict the evolving relationship between negative affect and craving during an attempt to quit smoking?" and "Which genes are most strongly associated with the time to smoking relapse?" The research team will use the new methods to examine new hypotheses, in order to advance the prevention of drug abuse and HIV infection.

Advancing Methods *(continued)*

Linda Collins

Sara Vasilenko and Stephanie Lanza

Engineering Better STI Prevention Using the Multiphase Optimization Strategy (MOST)

NIAAA Research Grant

Scientists need to design behavioral and biobehavioral interventions to be as effective, efficient, economical, and scalable as possible. MOST is a framework for engineering interventions that meet these criteria. MOST maximizes scientific knowledge while making the most efficient use of available resources.

Methodology Center Director Linda Collins, one of the original developers of MOST, has formed a partnership with a group led by David Wyrick, associate professor of public health education at the University of North Carolina at Greensboro. With a new R01 grant from the National Institute on Alcohol Abuse and Alcoholism (NIAAA), they will develop and pilot test an online intervention targeting the link between alcohol use and sexual risk behaviors among college students. The researchers will use the MOST framework to optimize the intervention. The resulting intervention will be designed to reduce heavy drinking, sexual risk behavior, and sexually transmitted infection (STI) among college students. It will also provide a knowledge base for STI- and HIV-risk interventions for other at-risk populations.

New Insights From A Long-Term Study: Applying New Methods for Intensive Longitudinal Data

NIDA Research Grant

Funded by the National Institute on Drug Abuse (NIDA), Monitoring the Future (MTF) is an ongoing study of Americans' behaviors, attitudes, and values from adolescence into adulthood. Based at the University of Michigan, MTF has provided many insights into substance use trends among America's youth while amassing a huge volume of data over its 40-year history. A new NIDA-funded R01 grant, "Dynamic Links Between Risk Factors, Substance Use, and Consequences: Ages 18-35," will apply innovative methods developed at The Methodology Center to examine the etiology of risky behavior in MTF's long-term panel data.

The project, which includes researchers from University of Michigan's Institute for Social Research and The Methodology Center, will apply time-varying effect modeling (TVEM) and other emerging techniques for the analysis of longitudinal data in order to understand how risk changes over an individual's life course. The researchers will study the extent to which young adult substance use leads to adult health consequences, such as substance use disorders, sexual risk behaviors, and physical/mental health problems. New methods for complex data are allowing scientists to gain new knowledge from existing panel studies like MTF.

Donna Coffman

Bethany Bray

Big Data Methods for mHealth

NIH BD2K Career Development Grant

One of the most exciting areas of the technology-driven expansion of complex data sets is mHealth, medicine delivered via mobile devices. A new grant is exploring mHealth data related to physical activity, diet, stress management, and substance use. In October 2014, Donna Coffman received a K01 award from the National Institutes of Health Big Data to Knowledge (BD2K) initiative. The grant provides training and research support for Donna's work on the analysis of mobile sensor data, genomics data, and ecological momentary assessments in order to promote healthy behavior. Donna's most recent research focused on propensity scores for causal inference; this K01 will provide training to study methods from computer science and informatics for the management and analysis of big data. Her mentors for the project include Runze Li, Verne M. Willaman Professor of Statistics at Penn State and Methodology Center principal investigator, Vasant Honavar, professor and Edward Frymoyer Chair of Information Sciences and Technology at Penn State, and Joshua Smyth, professor of biobehavioral health and medicine at Penn State. The methods developed in this project will be used to construct adaptive, individualized health behavior interventions delivered in real-time, real-world contexts.

Understanding Gambling as an Addictive Disorder

NCRG Seed Grant

In the fifth edition of the *Diagnostic and Statistical Manual of Mental Disorders,* the new, expanded definition of "addictive disorders" includes both gambling and substance use for the first time. A new grant from the National Center for Responsible Gaming (NCRG) will allow Bethany Bray to integrate and expand methods for latent class analysis and causal inference in order to examine how adolescent gambling and polysubstance use are linked to negative consequences in adulthood. By investigating how gambling functions as an addictive behavior, this research will improve understanding of how to create more effective prevention and treatment programs and will allow Bethany to develop a line of research responsive to the mission of the Collaborative Research on Addiction at NIH (CRAN) initiative.

Moving Forward

Aside from the work described here, other new research projects are arriving, including a NIDA-funded grant that will apply time-varying effect modeling to understand the epidemiology of drug abuse. All of the recent changes have given us a strong sense of momentum. We look forward to working with leading behavioral researchers to apply our new methods to emerging types of data. Working together, the scientific community can improve understanding of harmful behaviors and how to manage them effectively and efficiently. The next five years are going to be great—stay tuned!

New Research, New Surroundings

Among other recent changes, our physical home is brand new. We recently moved back onto Penn State's University Park campus and into the newly constructed Health and Human Development Building. A floating staircase anchors the building in the open atrium that welcomes visitors to the bright, modern space. The new conference room boasts windows on three walls and a view of downtown State College. After ten years we are very happy to be back on campus.

OUR FUNDING

Research at The Methodology Center is funded by grants from the National Institutes of Health, primarily the National Institute on Drug Abuse, the National Cancer Institute, the National Institute of Diabetes and Digestive and Kidney Diseases, and the National Institute on Alcohol Abuse and Alcoholism. The Methodology Center also receives significant support from Penn State's College of Health and Human Development.

TO LEARN MORE ABOUT THE METHODOLOGY CENTER, VISIT **methodology.psu.edu**