



Science of Behavior Change Common Fund Program Fourth Annual Meeting of Investigators

June 23-24, 2014
Bethesda, Maryland

Revised October 15, 2014

This summary report was prepared by Chandra Keller-Allen, Rose Li and Associates, Inc., under contract to the National Institutes of Health (HHSN271201300569P Requisition no. 3122806). The statements, conclusions, and recommendations contained in this document reflect both individual and collective opinions of the meeting participants and are not intended to represent the official position of the National Institutes of Health or the U.S. Department of Health and Human Services. We gratefully acknowledge review of and comments on a draft of this report provided by Mustafa al’Absi, Meike Bartels, Courtney Byrd-Williams, Barbara Fredrickson, Ian Gotlib, Scott Halpern, Johannes Haushofer, Richard Heyman, James Hudziak, Jonathan W. King, Santosh Kumar, Robert Levenson, Rose Maria Li, Julie Lumeng, Alison Miller, Megan Moreno, Mary Ellen Perry, Melissa Riddle, Kevin Ochsner, Michael Otto, Henry Saffer, George Slavich, and Timothy Strauman.

TABLE OF CONTENTS

| | |
|---|----|
| TABLE OF CONTENTS | 2 |
| EXECUTIVE SUMMARY | 3 |
| MEETING REPORT..... | 7 |
| Introduction | 7 |
| Current Perspectives on Behavior Change: Insights from the SOBC Expert Review Panel | 8 |
| The Future of SOBC..... | 11 |
| Grantee Updates: Session 1 | 16 |
| Self-Regulation Measures and Targets | 21 |
| Grantee Updates: Session 2 | 26 |
| Stress and Stress Reactivity Measures and Targets..... | 31 |
| Grantee Updates: Session 3 | 40 |
| Interpersonal and Social Processes Measures and Targets..... | 44 |
| Day 1 Summary Group Discussion..... | 51 |
| Creative Destruction of the Science of Behavior Change..... | 53 |
| Grantee Updates: Session 4 | 55 |
| Environmental Measures as Moderators..... | 60 |
| Closing Group Discussion | 62 |
| APPENDIX A: MEETING AGENDA..... | 66 |
| APPENDIX B: PARTICIPANT LIST | 68 |

EXECUTIVE SUMMARY

The goal of the National Institutes of Health (NIH) Common Fund's Science of Behavior Change (SOBC) Program has been to enhance understanding of the basic mechanisms of behavior change across a broad range of health-related behaviors, and, in doing so, unite disparate research fields and bridge basic and clinical science, ultimately leading to the development of more efficacious and efficient interventions. The SOBC Program has spent the past year evaluating progress made, identifying challenges yet to be addressed, and exploring logical next steps. This year's Annual Meeting of Investigators served a dual purpose—to learn of progress made on SOBC-funded research and to leverage the knowledge, experience, and expertise of SOBC-funded researchers and invited guests to inform opportunities for the future of the SOBC Program. The SOBC 2 Program proposes to explore behavioral targets and constructs in the broad domains of **self-regulation**, **stress and stress reactivity**, and **interpersonal and social processes** more deeply in the future.

SOBC 2 Target Classes

Self-Regulation

The nomological network in which the concepts of self-regulation and self-control reside is broad and includes constructs that figure centrally as hypothesized mechanisms, targets, or behavioral phenotypes in research on health behavior and behavior change across a wide range of conditions and developmental phases. Constructs in the target class of self-regulation appear to be indexing multiple mechanisms and processes, some likely distinct, others overlapping, and for which the developmental trajectory is not fully mapped out. This suggests—at a minimum—the need for more *cross-validation* to confirm findings by repeating experimental manipulations in one project using an independent assay technique from other research, as well as *cross-calibration*, to permit comparisons across projects where self-regulation/self-control has been measured differently, and where re-assessment is not possible. The goal of such work would be to determine the extent to which various measures of self-control and self-regulation are tapping distinct or overlapping mechanisms, and whether measures are performing similarly across populations, laboratories, and age groups. Work is also needed to determine which measures are appropriate for which contexts, which are redundant, and which truly assess targets that are engaged by interventions in ways that are meaningfully related to behavior change.

Stress and Stress Reactivity

Both chronic and acute stress exposures have been associated with maladaptive health behavior profiles. Heightened sensitivity to threat may constitute a stress-reactive phenotype that predisposes individuals to maladaptive behavioral and psychological responses when confronted with stressors. Although in many situations stress is closely related to processes involved in self-regulation, researchers have noted that the stress response system may have its own set of targets and downstream behavioral and physiological consequences—which suggests additional potential intervention and moderation strategies. Yet current measures of

stress exposures and stress reactivity (e.g., inventories of stress exposures, diurnal cortisol rhythms) have been inconvenient or difficult for behavior change researchers to include in actual trials. Future technological advances in stress measurement for observational and clinical studies hold promise for inclusion in behavior change research.

Interpersonal and Social Processes

The influence of social partners and social network members on health behaviors is a topic of substantial research, and numerous mechanisms have been identified that account for social influences on individual behavior. There are powerful social variables that can support or undermine health behaviors (e.g., attachment alliance, social grouping for safety, loneliness/social connectedness). People are sensitive to the social signal in a message and the social value of behavior. In addition to assays of individual consequences of interpersonal mechanisms (e.g., self-reports, physiological or neural responses in individuals), thorough assessment of interpersonal mechanisms requires systematic observational coding of putative causal processes.

Environmental Factors as Moderators

Prominent among the approaches to behavior change that have targeted the environment are those based on *behavioral economics*. Behavioral economics involves the study of the psychological, social, cognitive, and emotional factors that drive choice, including choice related to health-related behaviors (e.g., eating unhealthful foods, engaging in sedentary versus physically active behaviors). Research has demonstrated that individuals' decisions are sensitive to cognitive biases, such as default bias, that may be difficult to intervene on directly (i.e., within the individual) but might be possible to exploit via an intervention targeting the environment (i.e., a manipulation of the "choice architecture" in which a particular decision or behavior is enacted). Discussion during this session explored how environmental manipulations can be leveraged to induce large-scale behavior change and possible approaches to examining the interaction between environmental factors and behavior change targets at the individual (e.g., self-regulation, stress reactivity) and interpersonal levels.

Major Themes

Top-Down and Bottom-Up

There is a deep chasm between significant advances in research on basic mechanisms and the development of behavior change interventions. Meeting participants discussed different approaches to validating mechanisms of behavior change. A top-down approach starts with a model system that works and tests putative mechanisms within the intervention (outcomes → constructs). A bottom-up approach focuses on hypothesized mechanisms of behavior change and tests whether they can be measured, targeted, manipulated, and actually impact the target behavior (constructs → outcomes). Meeting participants had different preferences for these approaches, which were informed by their own expertise and experience. However, it was agreed that both approaches are valid and necessary to move the field forward and that existing theory can be a guide to bridging the two approaches. Regardless of approach, mechanisms need to be pursued in the context of a specific health problem.

Using an approach similar to the Research Domain Criteria (RDoC) project, SOBC 2 could support efforts to reach agreement on a set of intermediate phenotypes for each target class and then test the relevance of each phenotype for behavior change.

Behavior Is Complex and Context-Dependent

Behavior change works for certain people in certain environments. This complexity is an opportunity to test models and refine constructs. All three of the discussed target classes—self-regulation, stress and stress reactivity, and interpersonal and social process—are primed for more complex investigation and greater refinement. The “2.0” of behavioral science is a personalized model that accounts for unique person-by-cue-by-strategy-by-context interactions.

Constructs and Measures

There was general understanding and agreement among participants that the broad target classes refer to many constructs in multiple domains and are not constructs themselves. Several participants stressed that constructs need to be identified and refined before identifying, validating, and cross-calibrating reliable measures of those constructs with a process similar to that used by RDoC.

Measures need to tap the core consensus constructs using agreed upon metrics. The best measures will be those with demonstrable links to outcomes (e.g., what aspects of stress influence the immune system). Although reliability of a measure is necessary, validity for a particular purpose is typically not sufficiently explored. The first question after “is the measure reliable?” should be “to predict what, for whom, and under what circumstances?” The answer to the validity question will vary depending on the purpose. Research focused on determining the validity of measures for specified purposes would be incredibly useful to the field. Basic researchers and interventionists working together can bring measures to scale in ongoing trials.

Conclusion

The NIH-wide behavioral intervention portfolio analysis conducted by the SOBC Working Group found that few projects incorporated tests of mechanisms, even when the mechanisms hypothesized to be responsible for behavior change were explicitly stated in the application. This does not mean that the intervention was not built on the assumptions of mechanisms or a thoughtful theoretical foundation. But what is *not* being routinely included is a *test* of the mechanism. It cannot be enough to say “I think the reason the intervention will work is X mechanism.” The SOBC 2 Program is proposed to spur a paradigm shift in NIH-funded behavior change research such that the majority of future applications instead say “I think the reason the intervention will work is X mechanism, and I am going to test whether we moved X and then whether that moved the outcome.”

The vision for the SOBC 2 Program is that at the end of another 5-year period the effort will have supported work to systematically study and test a sample of promising intermediate-level

mechanisms along the pathway to multiple behaviors that impact a broader range of clinical endpoints and, in doing so, provide a proof-of-concept model for incorporating the testing of mechanisms into ongoing trials and produce validated and cross-calibrated measures for the constructs involved. A long-term goal is to change the behavior of the NIH and investigators by setting up the expectation that behavior change research will include the testing of mechanisms by providing measures, methods, and a model for doing so.

MEETING REPORT

Introduction

The goal of the National Institutes of Health (NIH) Common Fund's Science of Behavior Change (SOBC) Program has been to enhance understanding of the basic mechanisms of behavior change across a broad range of health-related behaviors, and, in doing so, unite disparate research fields and bridge basic and clinical science, ultimately leading to the development of more efficacious and efficient interventions. The SOBC Program has spent the past year evaluating progress made, identifying challenges yet to be addressed, and exploring logical next steps. This year's meeting served a dual purpose—to learn of progress made on SOBC-funded research and to leverage the knowledge, experience, and expertise of SOBC-funded researchers and invited guests to inform opportunities for the future of the SOBC Program.

Following a keynote address from SOBC Expert Review Panel member Michael W. Otto and an overview of the SOBC Program from Working Group Coordinator Jonathan W. King, current SOBC-funded investigators gave brief research updates highlighting activity in the past year, focusing on significant scientific advances. Presentations from invited speakers informed rich discussions on the areas that the SOBC Program intends to explore more deeply in the future: identifying behavioral targets and constructs in the broad domains of **self-regulation, stress and stress reactivity**, and **interpersonal and social processes**. Several discussion sessions were guided by three motivating questions:

1. What are the linkages, synergies, and commonalities across SOBC-funded projects and other behavioral research in how constructs in these domains are defined and measured?
2. What are the discrepancies or differences across SOBC-funded and other research in how constructs in these domains are defined and measured?
3. What are emerging approaches for defining and measuring common constructs in these domains, and what next steps are needed to unify the field?

The remainder of this report presents summaries of the scientific presentations and group discussions. The agenda is included as Appendix 1, and the list of participants is included as Appendix 2.

Current Perspectives on Behavior Change: Insights from the SOBC Expert Review Panel

Michael W. Otto, PhD, Boston University

Otto served as a member of the SOBC Expert Review Panel in early 2014 convened to evaluate the current program and provide recommendations for future directions.¹ The Expert Review Panel strongly supported the SOBC research agenda to coordinate efforts and allocate significant total resources toward addressing the challenges of health-related behavior change and maintenance. The Panel members agreed unanimously that the SOBC Program should not only be continued, but also expanded in scope and budget in a second 5-year term of support from the Common Fund. Without strategic leadership from the SOBC Program, research on the basic mechanisms of behavior change and its translation to clinical work will continue to be diffuse, disorder-specific, and disproportionately underfunded compared to its impact on the nation's health.

Behavior Change Research Is Critical for Realizing Benefits of Biomedical Research

Human behavior has been found to account for nearly 40 percent of the risk associated with preventable premature deaths in the United States.² Yet resources for behavioral and social research aimed at understanding basic mechanisms of behavior change across a broad range of health-related behaviors and the development and translation of effective and efficacious behavioral interventions remain disproportionately small compared to the potential return on investment in the form of dollars and lives saved. Behavior—such as screening, decision making, and adherence—is the critical variable that can increase or decrease the likelihood that the benefits of NIH-funded biomedical advances in treatments and interventions are fully realized.

Medication adherence is an example of how behavior change and maintenance is critical to realizing the advances of drug therapy. Drug development is well supported, yet resources of a similar magnitude are not allocated to understand the mechanisms of adherence. Maddox et al. found that 52 percent of patients stop taking their antidepressant medication within 12 weeks of initiating therapy.³ The strict type of adherence necessary for human immunodeficiency virus (HIV) medication is another example of how behavior is critical for patients to benefit from drug therapy. Mediocre adherence to HIV medication is a risk factor for medication resistance and

¹ The SOBC Expert Review Panel was chaired by Warren K. Bickel, Virginia Tech, and members in addition to Dr. Otto included John T. Cacioppo, University of Chicago; Linda M. Collins, Pennsylvania State University; Elissa S. Epel, University of California, San Francisco; Susan Fiske, Princeton University; Margaret Grey, Yale University; David Labison, Harvard University; and Rajita Sinha, Yale University.

² Schroeder, S. A. (2007). We can do better—Improving the health of the American people. *New England Journal of Medicine*, 357, 1221-1228.

³ Maddox, J. C., Levi, M., & Thompson, C. (1994). The compliance with antidepressants in general practice. *Journal of Psychopharmacology*, 8, 48-53.

death, a problem that might apply to some cancer drugs as well. Adherence is still an issue even with significant advances in the tolerability of HIV medication, particularly for populations struggling with multiple problems, such as HIV, drug addiction, and depression. The Life-Steps cognitive behavioral therapy for depression and adherence intervention involving understanding, motivational framing, and behavioral reversals was shown to be effective in a randomized controlled trial (RCT) for a population of HIV-infected injection drug users.⁴

Exercise adherence is another example. Science has shown that interval training is a tremendously effective strategy for burning fat and building muscle. However, intense interval training is so difficult to maintain that adherence to such exercises is low. People are more likely to become discouraged from such difficult routines and exercise less. Advances in biology need to be married to advances in behavior change to be effective.

The SOBC Program brought together basic and clinical scientists from multiple domains and disciplines to promote communication, common vocabularies, shared frameworks, and methods of comparison for diffusion of ideas across disciplines. This type of work is difficult but crucial. The many innovations in health care treatment and interventions—new technologies, procedures, drugs—must be applied through the mechanisms of behavior change to get patients and clinicians to take advantage of the biomedical innovations that will ultimately lead to the desired health outcomes at a population level.

Selected Recommendations

Funding Mechanisms

The SOBC Expert Review Panel was overall very impressed with the quality of the funded R01s and their contribution to advancing understanding of mechanisms of behavior change in various domains at different levels of analysis. The Panel also expressed firm support for the use of administrative supplements as a cost-efficient strategy for supporting the infusion of behavior change mechanisms into ongoing studies. Administrative supplements enable the SOBC Program to leverage the study infrastructure funded by the parent grant, infuse mechanistic research into a diverse set of ongoing clinical studies, and potentially shape the behavior of investigators. The Panel highly recommended using the administrative supplement strategy again in a potential future SOBC 2.

Cooperative agreements are another funding mechanism to consider for the future. Cooperative agreements would allow for increased coordination across a set of projects working toward the SOBC research agenda and for National Institutes of Health (NIH) project scientists to identify opportunities to provide input and guide investigators in infusing SOBC principles into their work and perhaps point them toward specific measures shown to be valid and reliable for a particular behavioral target. Otto referenced a recent Funding Opportunity

⁴ Safren, S. A., O’Cleirigh, C. M., Bullis, J. R., Otto, M. W., Stein, M. D., & Pollack, M. H. (2012). Cognitive behavioral therapy for adherence and depression (CBT-AD) in HIV-infected injection drug users: A randomized controlled trial. *Journal of Consulting and Clinical Psychology, 80*, 404-415.

Announcement (FOA) from the National Cancer Institute (NCI) as an example: *Fundamental Mechanisms of Affective and Decisional Processes in Cancer Control* (U01).⁵

Clearinghouse for Information

The SOBC Expert Review Panel also suggested that the group attend to visibility and dissemination to maximize the integration of its work into clinical science. One proposal, of which Otto was a vocal proponent, was to develop a publicly accessible clearinghouse of information on rate-limiting steps for behavior change. A public website cataloging studies, assessment targets, effect sizes, and findings could provide investigators with a tool to determine what to include in future studies and which measures work best under what conditions. The resource would support cross-validation by repeating experimental manipulations in one project using an independent assay technique from other projects and cross-calibration to permit comparisons across projects where a particular target has been measured differently, and where re-assessment is not possible.

Behavior Change Examples

Motivation

Our cultural notion of motivation is as an inward reservoir. People reference their level of motivation as if the tank were either empty or full. Instead, motivation can be conceptualized as a hierarchy of competing motivations, and which one wins depends on both external and internal contexts. Muraven and Baumeister and others conceived of self-regulation as a resource that can be depleted.⁶ Exercising self-control throughout the day—skipping French fries at lunch, coping with stress at work, dealing with negative affect, paying bills at night—fatigues the “muscle” of self-control and is influenced by stress. One way to conserve the muscle is by using chaining to move to a context that is more conducive to the desired behavior change. For example, getting up from the couch to run after self-control has been depleted seems like an insurmountable task. An easier strategy is to think in terms of smaller steps: get dressed for running, go outside, walk, run. Each step places the individual in a new motivational context, and, because of that context, the self-control push to the next step is easier to achieve.

Another example of using the environment to motivate behavior change is the use of fly etchings inside urinals, which has successfully reduced spillage by 85 percent according to industry studies. This is more effective than simply placing a sign from the management above the urinal that reminds men to “aim well.” Not only does the sign draw the person’s attention away from where you want them to aim, but also it relies on the person tapping into the motivation of others (i.e., the management wants me to do this) to induce behavior change. Interventions do not need to rely only on tapping an internal desire to change, but can capitalize on creating conditions in which a person is more likely to change his or her behavior automatically or in conjunction with other, already potentiated, motivations.

⁵ See PAR-14-067 at <http://grants.nih.gov/grants/guide/pa-files/PAR-14-067.html>.

⁶ Muraven, M. & Baumeister, R. F. (2000). Self-regulation and depletion of limited resources: Does self-control resemble a muscle? *Psychological Bulletin*, 126, 147-259.

Delay Discounting

There are individuals who seem naturally successful at behavior change. Part of the researcher's job is to identify characteristics of individuals for whom behavior change is more difficult and therefore might predict failure at a given behavior change intervention. Delay discounting—the devaluation of outcomes that occur more distally—is a good example of a characteristic that predicts the ability to change certain behavior, in particular, health behavior outcomes where a person must forgo current temptations to achieve longer-term beneficial outcomes. Working memory—in part determining the ability to hold two concepts in mind at once—appears to be important for a person's capacity for delaying immediate rewards in favor of a future goal and has been conceptualized as the ability to recall the future at points of temptation.

Distress Intolerance

Distress intolerance predicts a plethora of behaviors and behavioral patterns including addictive drug use;⁷ coping motives for substance use; illicit drug use patterns; degree of eating in response to negative affective episodes; avoidance of exercise and negative affect during exercise; sensitivity to pain and disability from dyspnea; sleep disruption; treatment drop out among inner city drug users; and general impulsive behavior. Distress intolerance serves to amplify stress and negative affect, leading to poor coping and various undesirable behaviors.

The Future of the SOBC Program

Jonathan W. King, PhD, National Institute on Aging

The SOBC Workgroup was formed in 2008, and funding for the program by the Common Fund began in fiscal year 2010. The program sought to confront the balkanization in the field of behavior change interventions—and in the basic science that supports them—and to create a unified science of behavior change to begin to tackle the substantial behavioral contributions to morbidity and mortality across a wide range of health and disease targets, central to the mission of multiple Institutes and Centers (ICs) across the NIH.

It is now widely appreciated that poor health behaviors (e.g., smoking, drinking, poor diet, lack of exercise, failures of adherence to medical regimens) account for a substantial proportion of disease burden in the United States. With data from McGinnis and Foege,⁸ Schroeder⁹ illustrated proportional contributions of a number of factors to premature deaths (those occurring before age 75), finding 40 percent of preventable deaths in the United States attributable to behavioral patterns. Although the final analysis on premature deaths attributable to behavioral risk factors based on more recent data examined by the Committee

⁷ Baker, T. B., Piper, M. E., McCarthy, D. E., Majeskie, M. R., & Fiore, M. C. (2004). Addiction motivation reformulated: An affective processing model of negative reinforcement. *Psychological Review*, *111*, 33-51.

⁸ McGinnis, J. M. & Foege, W. H. (1993). Actual causes of death in the United States. *Journal of the American Medical Association*, *270*, 2207-2212.

⁹ Schroeder, 2007.

on Population Meeting on Premature Death at the National Academy of Sciences (NAS) is not yet complete, it is clear that the burden attributable to behavioral factors has, if anything, increased over the past 20 years.¹⁰ The scope of disease burden, health care costs, and preventable deaths is an enormous public health challenge. Despite widespread awareness on the part of clinicians that improvements in health behaviors are essential for health promotion, prevention, and disease management, it remains exceptionally difficult to motivate people to initiate and maintain behavioral change. Can it be done? And what is the evidence that if done, it will have substantial clinical impact?

One notable success in the area of behavior change is the Diabetes Prevention Program study funded by the National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK).¹¹ The study compared drug therapy (metformin), lifestyle change, and a control in pre-diabetic individuals. A 7 percent weight reduction and 2.5 hour per week activity increase led to a 58 percent reduction in the cumulative incidence of Type 2 diabetes in insulin-resistant individuals. The lifestyle advantage holds up 10 years later.¹² Although it is clear that the lifestyle group has maintained its advantage over the placebo group, the exact mechanism(s) of behavior change and differences among subjects in response to treatment are unclear.

Over time, it has become apparent that at least three different scientific divides have thwarted development of a more unified science of behavior change. First, lessons from basic science—including newly emerging mechanistic insights from transdisciplinary domains of behavioral science—are rarely applied to the search for intervention targets for behavior change. Second, although NIH supports much basic work on behavioral and biobehavioral mechanisms, every problem behavior has its own clinical endpoint, sometimes its own IC, and over time its own research community, such that even researchers who study closely related endpoints or problem behaviors have worked independently of each other, limiting the prospects for generalization in clinical science. Finally, there remains an artificial separation between basic science that identifies mechanistic intervention targets and clinical science that seeks to modify those targets to affect clinical endpoints. The SOBC Program has sought to bridge all three divides since its inception.

SOBC and SOBC 2

The SOBC Program goal has been to capitalize on emerging basic science to accelerate the study of common mechanisms of behavior change. The program supported lab and field studies to delineate what the common behavior change mechanisms are and how they are engaged in different contexts. The SOBC Program scientific meetings have helped to break down

¹⁰ U.S. Burden of Disease Collaborators. (2013). The state of U.S. health, 1990-2010: Burden of diseases, injuries, and risk factors. *Journal of the American Medical Association*, 310, 591-606.

¹¹ <http://diabetes.niddk.nih.gov/dm/pubs/preventionprogram/>.

¹² Davis, N. J., et al. (2013). Predictors of sustained reduction in energy and fat intake in the Diabetes Prevention Program Outcomes Study intensive lifestyle intervention. *Journal of the Academy of Nutrition and Dietetics*, 113, 1455-1464.

disciplinary boundaries, begin collaborations, and expand perspectives on the mechanisms of behavior and behavior change in the laboratory and the field.

It has been proposed that SOBC 2 would implement an intervention development pipeline (i.e., the experimental medicine approach) for behavior change interventions similar to that used for the development of drugs and devices (see Figure 1). For drug development, there is a familiar sequence of steps from target identification through lead optimization prior to preclinical studies in animal models, obtaining an Investigational New Drug (IND) application from the Food and Drug Administration (FDA), and the sequence of Stage I, II, and III clinical trials, followed by FDA approval and then (hopefully) by additional trials including effectiveness trials in the real world.

The behavior change intervention development process does not operate similarly. There is often diffuse preclinical activity; preclinical studies in animal models are uncommon, but small clinical trials in people are common. There is no IND step for behavioral interventions, and although there can be a sequence of clinical trials ending in something analogous to a Stage III clinical trial, such as the Diabetes Prevention Program, even if it is successful, the follow-up is likely to scale up the intervention and make it less expensive. However, because the efficacious intervention is often complex, but untested in any of its specifics, the “active ingredient” is not well understood and the scaled-up version might not be efficacious.

SOBC 2 is intended to support work that will replace the current unsatisfactory situation with a sequence of steps based on the experimental medicine approach, but with the acknowledgment that the preclinical studies will often provide feedback relevant for additional target identification or validation, as well as the possibility to test whether interventions can be optimized either in general or for some subpopulation. Future clinical trials will at least give researchers not only (hopefully) improved interventions, but also the assays needed to assure that the putative targets of an intervention are engaged, or the ability to optimize or personalize the intervention on the basis of the assays performed.

Intervention Development Pipelines

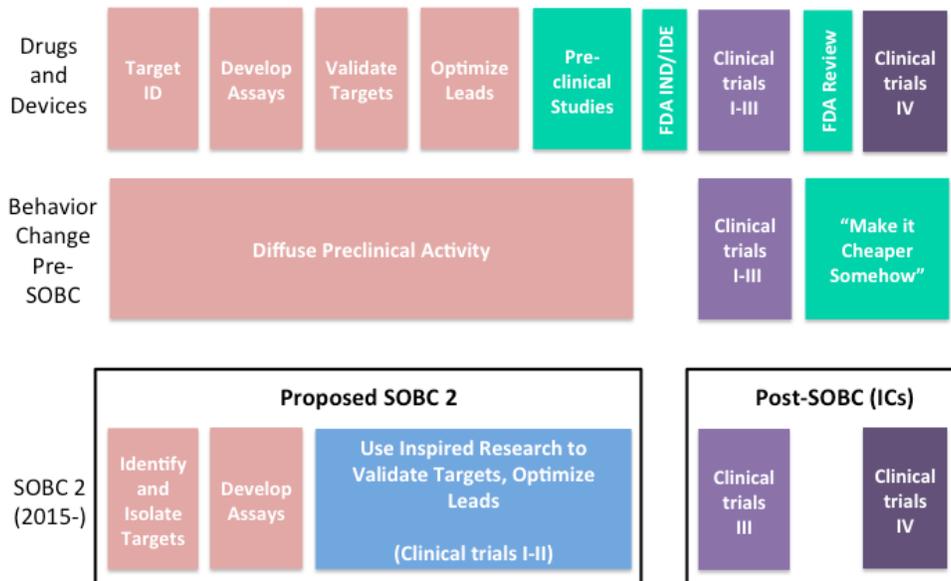


Figure 1: Drugs and devices and SOBC 2 intervention development pipelines.

It is envisioned that projects funded under SOBC 2 will work to

- *Identify the most promising intervention targets whose engagement drives behavior change.* Projects will take broad concepts related to behavior and behavior change (e.g., self-control or self-regulation) and form consensus on what specific processes are involved and how they are related to each other, so potential targets for intervention can be isolated.
- *Develop appropriate assays to measure target engagement.* With better-defined targets, specific assays that are sensitive to change can be developed and therefore be reliable indicators of target engagement. For example, through his work on manipulating attentional bias away from negative stimuli, Ian Gotlib also knew the degree to which a participant was or was not attentionally biased by his or her performance on a specific task, and that would be a candidate assay. Similarly, if the response to acute or chronic stress is hypothesized to involve maladaptive behaviors, then a behavioral or physiological measure of stress would be quite valuable.
- *Validate putative targets in laboratory and clinical contexts.* Because behavior is quite complex, it is possible that some promising putative targets for which a convenient assay exists turn out not to be valid targets in the sense that engaging them with an intervention and changing them in the way predicted to affect the outcome does not appear to work. But, by the same token, some targets will be involved in multiple forms of behavior (often in the same person); therefore, separate targets for every specific kind of maladaptive behavior might not be needed.

- *Systematically improve behavioral trial designs to incorporate measures of target engagement throughout the intervention-target-clinical endpoint pathway.* SOBC 2 will succeed if at the end of the 5-year period, ICs are in the position to improve the efficacy and effectiveness of behavior change interventions.

NIH Portfolio Analysis

As part of the review of the SOBC Program and to identify gaps that could be addressed by SOBC 2, Working Group members conducted a portfolio analysis of NIH-funded studies that test mechanisms of behavior change from a selection of ICs. The analysis demonstrated that among development and pilot testing studies, which comprised 44 percent of the studies in the analysis, 54 percent of the studies did not include tests of mechanisms. Among efficacy testing studies in the analysis, 37 percent did not include tests of mechanisms and 34 percent included exploratory testing of mechanisms of behavior change. A great deal of resources are currently invested in behavioral trials, and the expense of increasing sample size yet only a minority of studies include tests of mechanisms to determine what it is about the intervention that is engaging a target and changing behavior.

Key Targets for Behavior Change

Discussion at multiple SOBC Program meetings, findings from grantee work, and the Expert Review Panel suggestions have yielded general agreement on four broad domains that include several specific targets that drive behavior relevant to multiple diseases and health outcomes:

- *Self-Regulation:* Manipulating attention (in place of exerting effortful control) can lead to behavior change.
- *Stress and Stress Reactivity:* Decision making under stress may ignore long-term consequences.
- *Interpersonal and Social Processes:* Social partners and social network members can influence behavior.

In addition, environmental and social factors can induce or suppress behavior (so the target can be the environment), and environment (or context) plays a role in the other target classes.

Rarely, if ever, does an interesting intervention directly cause a behavior change. More frequently there is a putative target in mind that is being manipulated with the hypothesis that this target is causally connected to the behavior we wish to change. But the target cannot be meaningfully tested unless there is a way of knowing whether the target can be hit. Assays that validly measure changes in target “activity” are needed to determine whether an intervention can successfully cause such a change in the target. These assays could take the form of behavioral tests, neuroimaging data, an endocrine assay such as cortisol or alpha amylase, or gene expression. Once assays are in place, it can be verified whether the manipulation engages

the target. For behavioral interventions, this is most often done within the context of an intervention trial involving the proposed intervention.

The Common Fund—at the time of this meeting—had not yet approved or denied the SOBC 2 proposal. Details of the program are thus subject to further refinement.

Grantee Updates: Session 1

Investigating Affective and Genomic Mediators of Sustained Behavior Change

Barbara Fredrickson, PhD, University of North Carolina at Chapel Hill

This project is examining the affective and genomic mediators of sustained behavior change and is framed by Fredrickson's Upward Spiral Theory of Lifestyle Change. The overarching hypothesis of this theory is that positive emotions create *nonconscious* and *increasing* motives for wellness behaviors. To the extent that positive emotions are elicited by a wellness behavior, the emotions come with the ability to create nonconscious wanting that draws a person to repeat the wellness behavior. This broadens awareness and builds resources within the person (i.e., the broaden-and-build theory of positive emotions).¹³ One finding thus far is that positive emotions build certain resources that serve to amplify the positive emotion yield of the wellness behavior—for example, cardiac vagal tone and psychological mediators.

The second study in the project is an RCT comparing loving kindness meditation (LKM) to mindfulness meditation (MM) as a way to test a wellness behavior with and without positive emotions. A particular objective of LKM is to increase the ability to generate positive emotions. Data collection for time 3—where measures of sustained behavior change were used—was completed in May 2014. Analyses of those data are ongoing.

Data collected from time 1 and time 2 indicate that compared to MM, LKM increases daily positive emotions from baseline to week 6 and the increase is found to be moderated by polymorphisms in oxytocin genes OXTR and CD38. Previous research shows that nonrisk carriers of OXTR are more likely to show pro-social behavior, and this study shows they are more responsive to LKM. This pattern is not seen in the risk carriers, and the same pattern holds for CD38.

In the first SOBC study, it was demonstrated that oxytocin as a person-variable predicted positive emotion yield of a wellness behavior. In her collaboration with Steve Cole, Fredrickson used a functional genomics approach to look at changes in gene expression, especially in an a priori area of interest—conserved transcriptional response to adversity (CTRA). CTRA is based on evidence of a 53-gene composite that reflects increased expression of pro-inflammatory genes, decreased expression of antiviral genes, and decreased expression of antibody synthesis

¹³ Fredrickson, B. L. (2013). Positive emotions broaden and build. *Advances in Experimental Social Psychology*, 47, 1-53.

genes.¹⁴ Data from this study indicate that LKM reduces adversity-related patterns of gene expression (See Figure 2). The data collected will enable the investigators to describe participants psychologically and biologically and to determine whether it is possible to triage them into the meditation approach that is most beneficial based on individual characteristics.

LKM reduces adversity-related patterns of gene expression (AIM 2)

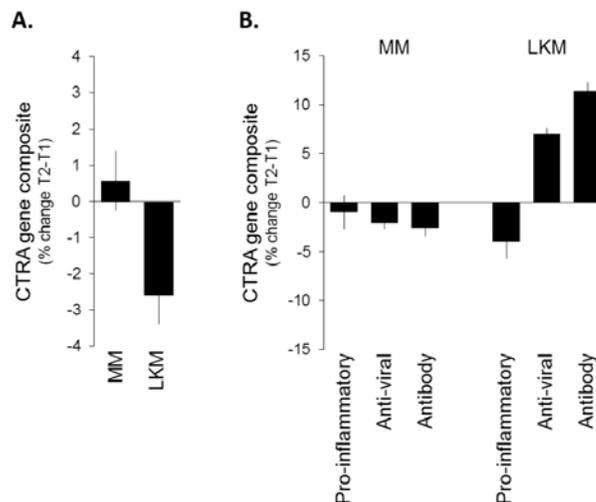


Figure 2: Loving kindness meditation reduces adversity-related patterns of gene expression.

Self-Regulation Failure: Identifying and Modifying a Risk Phenotype

Timothy Strauman, PhD, Duke University

This project was designed to test a trans-diagnostic model of a risk phenotype for self-regulation failure, as well as cognitive-behavioral micro-interventions targeting the mechanisms underlying the phenotype. The phenotype is an interaction among (1) individual differences in self-regulatory style (regulatory focus), (2) catechol-*O*-methyl transferase (COMT) genotype, and (3) chronic goal failure. This presentation focused on cortical activation “hot spots,” behavioral evidence for dysfunction, and evidence for the efficacy of theory-based micro-interventions.

Imaging data show that there is robust left and occipital prefrontal activation when people are encouraged to think about what kind of person they want to be, and this activation is correlated with individual differences measures. Activation at the left occipital prefrontal cortex site is specific to promotion priming and is correlated with self-reported strength of promotion orientation. However, individuals who met behavioral and genetic criteria for the hypothetical

¹⁴ Cole, S. W. (2009). Social regulation of human gene expression. *Current Directions in Psychological Science*, 18, 132-137.

Fredrickson, B. L., et al. (2013). A functional genomic perspective on human well-being. *PNAS*, 110, 13684-13689.

risk phenotype showed increasing activation across the block of promotion trials. The risk phenotype group showed greater activation whether goals happened to be goals they thought they were meeting or goals they thought they were not meeting. Those without the risk phenotype responded to a relevant goal and then moved on.

A well-validated goal-switching task was used to observe predicted dysfunction within the risk phenotype group when subjects were reminded of failure experiences. The goal-switching task allowed for observation of relative flexibility versus perseverance, and it added promotion and prevention (and neutral) priming. It was more difficult for participants with met/met typed COMT genes to disengage from a goal when they received failure feedback. This finding suggests that the process could be observed as it happens.

One goal of the grant was to pilot micro-interventions applicable to the real world that target dysfunctional self-regulation. Regulatory focus theory suggests that the risk phenotype group would benefit from interventions that reduce promotion system engagement when the individual is encountering repeated promotion goal failure feedback that is not responsive to additional effort. One strategy drawn from social cognition research is to have the individual engage in counterfactual thinking (“What might I have done differently?”). However, which type of counterfactual to use depends on whether the individual is engaged in promotion or prevention goal pursuit. Findings indicate that counterfactuals targeting promotion/prevention selectively reduce distress in the risk phenotype group. Additive counterfactuals (e.g., what could I do in the future) led to decreased anxiety, and subtractive counterfactuals (e.g., what mistake did I make) led to decreased sadness. The next steps are to look at gene expression and the ability to identify the risk phenotype without having to genotype individuals.

Mechanisms Underlying Attentional Bias Training in Children at Risk for Depression

Ian H. Gotlib, PhD, Stanford University

Gotlib and colleagues are exploring interventions to prevent the onset of psychopathology in young children at risk because of family history. An earlier study showed that like depressed adults, never-disordered daughters of mothers with recurrent depression exhibit negative attentional biases and a lack of positive biases, suggesting a causal role in major depressive disorder (MDD).¹⁵ Inferences can be made about attentional bias based on a task that tracks participants’ attention when they are shown negative, neutral, or positive faces.

Recent research has documented the beneficial effects of Attention Bias Training (ABT) in reducing symptoms of anxiety and depression. Gotlib is recruiting 50 girls at familial risk for MDD and using ABT (and a sham control condition) to teach them to decrease attention to negative stimuli and increase attention to positive stimuli. The SOBC Program administrative supplement is supporting work to assess the neural mechanisms underlying the effectiveness of ABT. The investigators are examining changes in patterns of neural activation related to

¹⁵ Joormann, J., Talbot, L., & Gotlib, I. H. (2007). Biased processing of emotional information in girls at risk for depression. *Journal of Abnormal Psychology, 116*, 135-143.

decreased attentional capture by negative stimuli, increased inhibition of negative material, and decreased negative self-referential thinking. Preliminary data document possible neural mechanisms underlying changes in information/attentional processing as a result of ABT in young girls at familial risk for depression.

Participants are scanned pre- and post-ABT to assess neural activations related to decreased attentional capture by negative stimuli, increased inhibition of negative material, and decreased negative self-referential thinking. The training consists of six sessions of real or sham ABT over 2 weeks. The same scans are repeated post-training. Preliminary data show that the 2-week ABT increases positive affect and reduces heart rate during a stress task by almost five beats per minute, which is a significant change for children.

Pre- and post-training functional magnetic resonance imaging (fMRI) scans of an emotion interference task shows training-based changes in activations to sad faces. The pre/post decrease in amygdala activation is correlated with change in attentional bias. There is also an increase in prefrontal cortex activity (indicating cognitive control activation) from pre- to post-ABT that is also strongly related to each child's change in attentional bias.

The increased attention to positive stimuli and decreased attention to negative stimuli as a result of ABT is associated with decreased amygdala reactivity to sad distractors and increased prefrontal reactivity to sad distractors. These findings are the first to demonstrate a direct relation between cognitive training and changes in brain function in children at risk for the development of psychopathology. Examining and testing mechanisms are necessary to understand how behavioral interventions produce their effects, which is a major objective of the SOBC Program, particularly as expressed in the administrative supplement FOA.¹⁶

Gotlib and colleagues are currently combining the obtained neural data with data concerning changes in behavioral and clinical functioning (both cross-sectional and longitudinal) to understand more precisely how ABT achieves its beneficial effects. This will set the stage for changing neural functioning more directly via neurofeedback. The neurofeedback intervention involves real and sham tasks in which participants are shown a thermometer that displays the degree of activation in their brain salience network in real time. Participants are asked to make the thermometer reading go down as they look at a negatively valenced picture. In the sham condition, participants are shown another girl's feedback from the real neurofeedback condition—they think they are reducing their own activation but they are not. The 12 girls run through the real neurofeedback condition so far subsequently reduced their level of stress in response to stimuli, as measured by cortisol.

¹⁶ PA-12-119 *Use-Oriented Basic Research: Change Mechanisms of Behavioral Social Interventions* can be found at <http://grants.nih.gov/grants/guide/pa-files/PA-12-119.html>.

The Development of Emotion Regulation Mechanisms Impacting Health

Kevin Ochsner, PhD, Columbia University

Ochsner's project is being conducted in collaboration with Walter Mischel, BJ Casey, and Jennifer Silvers. The premise of this project is that children and adolescents are faced with affective triggers in everyday life (e.g., rejection) and some are repetitive. Individuals will vary in their reactions to these cues and over time; repeated exposures to them can lead to maladaptive health or mental health outcomes. Ochsner hypothesized that individuals would also vary in top-down regulatory control. For individuals and across development, this project seeks to identify behavioral and neural signatures of reactivity and regulation that predict vulnerability to maladaptive health outcomes.

The first phase of the project involved recruitment of about 200 individuals aged 6 to 18 years for assessments of behavior (e.g., ability to delay gratification, physical and mental health) and the brain (e.g., fMRI for affective reactivity, cognitive reappraisal, attentional control, and structural MRI for gray and white matter). These measures were repeated 2 years later. Only about two-thirds of the participants were "scannable," meaning many of them simply could not remain still enough to obtain useable scans.

Scans of 105 participants indicate that younger children have a higher baseline of craving, but they can effectively regulate when told to imagine the desirable treat is far away and they are looking at it as if through a camera lens.¹⁷ When responding naturally to images of tempting foods, older individuals showed increased activation of the dorsolateral and ventrolateral prefrontal regions associated with cognitive control as well as decreased reactivity in the ventral striatum, a region associated with reward. These data suggest that weaker baseline craving for older individuals may reflect a combination of weaker initial reactivity and the deployment of regulatory mechanisms even when participants are not instructed to do so.

When regulating, the results also showed interactions between cognitive control systems that support regulation and affect systems that trigger emotional responses. Here, all participants showed increased activity in regions associated with cognitive control, including the dorsal anterior cingulate cortex and the dorsolateral, medial, and ventrolateral prefrontal cortices—accompanied by decreased activity in a region commonly associated with reward learning, the ventromedial prefrontal cortex. Together, these data suggest that younger children can regulate when instructed to do so, but only adolescents and young adults regulate spontaneously, which leads to reports of weaker baseline craving for older individuals.

Another critical finding related the ability to delay gratification, as assessed using Mischel's classic task, to the ability to regulate craving, as assessed in the scanner. In a sample of 6- to 13-year-olds two key findings were obtained. First, replicating prior work, delay ability predicted body mass index (BMI) percentile—in this sample high-delay kids had lower BMIs compared to

¹⁷ Silvers et al., (under revision). *Psychological Science*.

low-delay kids.¹⁸ Second, the tendency to recruit left ventrolateral prefrontal cortex in the regulation of craving task (in an entirely separate testing session a month later) fully mediated the relationship between delay ability and BMI. This suggests that delay ability—and its ability to predict meaningful health outcomes—depends critically on prefrontal control systems engaged in the regulation of craving task.

Finally, Ochsner reported interesting preliminary structural data suggesting that high-delay children have thicker regions of ventrolateral and ventromedial prefrontal cortices. He is now working to identify profiles across all the regulatory tasks that might leave children most vulnerable for various health outcomes.

Self-Regulation Measures and Targets

Facilitators: William Klein, PhD, National Cancer Institute; Lisbeth Nielsen, PhD, National Institute on Aging; Kevin Ochsner, PhD, Columbia University; Timothy Strauman, PhD, Duke University

The SOBC Working Group drew on the collective experience and expertise of the grantees and invited presenters to reflect on the SOBC 2 proposal framework, provide input on future directions, and identify critical features, considerations, or processes for moving the research agenda forward. Self-regulation is a broad grouping of multiple constructs and frameworks. There are many self-regulation processes in a number of domains that have been identified as central to behavior change.

Based on the portfolio analysis of NIH-funded studies, it appears that many projects do not include measures of the relevant mechanism, even though they are motivated by a hypothesis that some element of self-regulation failure is the reason for a problem behavior (e.g., overeating, substance use). SOBC 2 seeks to support work that can define and refine a subset of promising self-regulation constructs; identify/develop, validate, and cross-calibrate measures of those constructs; and then support the systematic use of those constructs and measures in ongoing behavior intervention research.

Models for SOBC 2 Approach and Activities

Research Domain Criteria

The National Institute of Mental Health (NIMH) launched the Research Domain Criteria project (RDoC) to support the development, for research purposes, of new ways of classifying psychopathology based on dimensions of observable behavior and neurobiological measures. The effort is to define basic dimensions of functioning (such as fear circuitry or working memory) to be studied across multiple units of analysis, from genes to neural circuits to behaviors, cutting across disorders as traditionally defined. The intent is to translate rapid progress in basic neurobiological and behavioral research to an improved integrative understanding of psychopathology and the development of new and/or optimally matched

¹⁸ Silvers et al. (in preparation).

treatments for mental disorders. NIMH anticipates that research grants employing this new experimental classification will represent an increasingly large share of its funding portfolio in coming years. Janine Simmons (NIMH), Gotlib, and several grantees have been involved in the RDoC effort. The process is intended to be open and transparent to the research community.¹⁹

RDoC Working Groups were guided by three criteria in the selection of constructs for the RDoC Matrix.²⁰ First, the inclusion of a construct was constrained by whether a particular brain circuit or area could reasonably be specified that implements that dimension of behavior. Given the complexity of the brain and of behavior, this was more ambiguous in some cases than others; some constructs, such as attention, reflect activity spread relatively diffusely over many brain areas, while attachment behavior may similarly reflect neurotransmitter and hormonal functions (e.g., oxytocin) acting at disparate locations throughout the brain. Second, an attempt was made to maintain a reasonable “grain size” that would permit a tractable listing of the major functional dimensions of behavior. Although it is recognized that there may be important and meaningful sub-constructs that could be considered (e.g., various types of aggression), an overly specified list could become unwieldy and excessively long. Third, the constructs are based on current literatures that have provided a neurobehavioral research base for each of the entries.

Gotlib remarked that the advantage of RDoC is that it has units of analysis for particular domains and constructs and the focus is mechanisms. Replicability is not explicitly part of RDoC, but of course it is very important.

Simmons noted that RDoC did not use “self-regulation” as a construct because it was deemed to be too broad. RDoC constructs need to have enough scientific evidence to support their link to specific circuits without overlapping with other distinct constructs. The original intent, similar to what SOBC 2 is proposing, was to demonstrate valid measurement, test-retest reliability, and inter-individual sensitivity of the constructs and then to use them as targets for intervention.

It is possible that a team process, similar to the RDoC Working Groups, that identifies, defines, and refines constructs (and related measures) that fall under the broad behavioral concept of self-regulation would work for what is being proposed for SOBC 2. Researchers from multiple fields, including psychometricians and methodologists, who bridge basic and clinical research would be tapped to ensure all the relevant expertise is incorporated into such an effort.

Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia

Ochsner participated in the Cognitive Neuroscience Treatment Research to Improve Cognition in Schizophrenia (CNTRICS) initiative²¹ and described its similarities to RDoC and its applicability as a model process for the work SOBC 2 proposes to do.

¹⁹ Additional information about RDoC can be found at <http://www.nimh.nih.gov/research-priorities/rdoc/index.shtml>.

²⁰ The RDoC Matrix can be found at http://www.nimh.nih.gov/research-priorities/rdoc/nimh-research-domain-criteria-rdoc.shtml#toc_matrix.

²¹ Additional information about CNTRICS can be found at <http://cntrics.ucdavis.edu>.

The primary goals of CNTRICS are (1) to identify a set of cognitive systems and component processes to be targeted for treatment development in schizophrenia; (2) to delineate and target the psychometric and pragmatic issues relevant to the development of tasks that measure the function of these cognitive systems, including (a) issues related to the measurement of specific cognitive processes in the face of generalized performance deficits; (b) the establishment of benchmarks for reliability and stability of experimental cognitive measures, and (c) identifying strategies to optimize these properties during the development of experimental cognitive measures; and (3) to develop specific measures of target cognitive processes that can be implemented as behavioral tasks as well as in noninvasive functional neuroimaging studies using techniques such as fMRI and electroencephalography (EEG) and event-related potentials (ERP).

Through a series of meetings and publications led by Cameron Carter (University of California, Davis) and Deanna Barch (Washington University), interdisciplinary groups of researchers selected the relevant cognitive constructs, identified psychometric issues with measuring those constructs, and, through consensus and careful documentation of selection criteria, identified the best measures for each construct. Investigators could then apply for funding to test the selected measures to determine whether they were appropriate for use in clinical trials. Ochsner's laboratory developed several of the measures that were the most successful. He deemed the process a valuable exercise for the field. CNTRICS II is now focusing on the evaluation of biomarkers of cognitive function.

Tapping Existing Resources and Data

Klein offered the Grid-Enabled Measures (GEM) Database supported by the National Cancer Institute (NCI)²² as an example of a dynamic collaborative tool that enables users to add constructs or measures to the database; contribute to and update existing information (metadata) about constructs and measures; rate and comment on measures to drive consensus on best measures; access and share harmonized data; and search for and download measures. GEM is an example of a collaborative citizen science approach that might serve as a model for the type of clearinghouse or registry Otto suggested.

Otto clarified that the registry the Expert Review Panel envisioned also includes the ability to compute effect sizes and modifiers and would require entry of characteristics such as environment, cue, context, and a description of how the measure worked or failed to provide emergent information on how measures of selected constructs are faring in a given situation. Grantees would be required to submit their measures, data, and related information to the registry. Such a resource would enable connections across researchers and laboratories to be made on a larger scale.

Ochsner noted that there are many opportunities for mining existing data sources and working with people with technical expertise to advance refinement of constructs and related

²² Additional information about GEM can be found at <https://www.gem-measures.org/Public/Home.aspx>.

measures. For example, the Affective Computing Group at the MIT Media Lab²³ has developed methods for crowdsourcing reappraisal, which is something that Ochsner's research group had discussed but did not have the technical expertise to operationalize.

Constructs

Participants discussed the importance of focusing on defining and refining the constructs that fall in the target class of self-regulation and *then* identifying, validating, and cross-calibrating reliable measures of those constructs. Gotlib, in particular, noted that a focus on constructs and the mechanisms that need to be measured, similar to what was done with RDoC, is critical for moving the field forward.

Throughout the discussion, there was general understanding and agreement among participants that the term "self-regulation" refers to a broad class of constructs in multiple domains and is not itself a single construct. Participants also agreed that clarity of language and common definitions are critical because different vocabularies and terms are used in different research fields to reference the same or similar ideas. Any effort to advance a unified science of behavior change needs to use clear and consistent language.

Measures

Robert Levenson noted that although reliability of a measure is necessary, validity for a particular purpose is typically not sufficiently explored. For example, if the purpose of a desired measure is to predict whether an individual will continue exercising once he/she starts (i.e., a measure of adherence), then different measures for this specific purpose should be evaluated (i.e., is a 5-item self report, ability to delay gratification task, or scan the most valid measure?). The first question after "is the measure reliable?" should be "to predict what, for whom, and under what circumstances?" The answer to the validity question will vary depending on the purpose. Support to build a body of research focused on determining the validity of measures for what purpose would be incredibly useful to the field.

Alison Miller added to this idea by noting the importance of defining self-regulation constructs within the context in which they are being measured. The validity of a measure might depend on the context in which it is used (e.g., a supportive setting in which a person is motivated or a stressful and challenging situation). For example, in her work with toddlers, co-regulation in parenting (e.g., how well a parent promotes self-calming in a challenging situation) is a significant factor.

Levenson also defined two types of measures of regulation: measures of ability and measures of performance. The ability to delay gratification test (i.e., the marshmallow test) is an example of a measure of performance (i.e., what do you do in a particular situation?). The researcher provides a framework (and in this case, a choice), but no explicit instructions, and then

²³ Additional information about the Affective Computing Group can be found at <http://affect.media.mit.edu>.

determines what the person actually does. On the other hand, performance measures (when people are told what to do and then their ability to do so is measured) are extremely useful in gauging capacities under particular conditions. But the two kinds of measures are not interchangeable. For example, a person who is masterful at using a reappraisal strategy to down-regulate emotional arousal when told to do so in the laboratory may be disinclined or unable to use this strategy in everyday life. Another problem is that there are many self-report measures of ability and performance whose relationship to actual behavior has not been adequately established (and that may, in reality, not be highly related).

Fredrickson agreed that individuals are not very good at self-reporting general tendencies. However, self-reports in the size at which the behavior occurs can be very accurate and useful. For example, the day reconstruction method allows us to capitalize on the witnessing of behaviors in smaller chunks of time, and it reduces the likelihood that a respondent provides a rosy glow of general perception in retrospect. Eileen Shinn added that performance measures in real-life situations should be linked to the individual's health state and desired health outcome.

Megan Moreno suggested that in addition to self-report strategies, there are novel measurement tools that take advantage of mobile technologies that can capture aspects of a patient's daily life as it happens and then transmit the information to a researcher or clinician. For example, texting could facilitate ecological momentary assessment (EMA) or Instagram could be used to document meals. There are many possibilities for using mobile technologies for documenting behavioral decisions in creative ways.

Context

Ochsner described his theory of "Emotion Regulation 2.0." Looking across studies, it is generally understood that emotion regulation involves a specific set of prefrontal systems and modulates the salience network. But this understanding does not reveal anything about the unique profiles of individuals. Some are sensitive to particular cues, some use strategies more easily than others, and some abilities may depend on context. The 2.0 of behavioral science is a personalized model that accounts for unique person-by-cue-by-strategy-by-context interactions. In the cognitive neuroscience field, performance and neural activity on one task does not predict performance and neural activity on another task even though it looks like the same system is engaged for both tasks. Part of this is a methodological problem—the imaging is not sensitive enough—but another part of this problem is conceptual. Most cognitive neuroscience measures have no test-retest reliability. Research that systematically tests whether the existing tasks and measures can reliably test the target across time and individuals is needed. In summary, the person-by-cue-by-strategy-by-context interactions are related to the test-retest reliability of measures, and these issues add a layer of complexity to the process.

Elizabeth Phelps wondered how much changing emotion regulation behavior depends on decisions within the individual versus changing circumstances that can alter the behavior of the individual, either knowingly or not. Nielsen added that in the area of conscientiousness and

healthy aging there is sometimes a tension between those who want to intervene on within-individual targets to alter noncognitive skills that make the person more suited to regulate behavior and those who want to modify the environment to achieve the same desired behavioral outcome. Behavior interventions could target the individual or the context or both, but what is important is that the target is precisely defined, engaged, and measured.

Ochsner offered two classes of strategies: those that can be implemented by the individual or those that can be implemented within the social environment. There are different variables that determine when reappraisal makes sense, but there are other contexts in which the environment helps to provide an appraisal frame from the outside. Some strategies can only be used in the environment (e.g., removing all junk food from public school settings), some strategies can only be used by the individual, and some strategies may be applicable to both (e.g., situation selection or modification).

Grantee Updates: Session 2

Emotion and Decision Making

Elizabeth Phelps, PhD, New York University

The two specific aims of this project address the questions (1) Can we use the tools of affective neuroscience and neuroeconomics to characterize more precisely how and when emotion is incorporated into the computation of subjective value? and (2) Can we use the tools of affective science to change emotion and change choice?

Arousal and Loss Aversion

Previous research has shown that people can regulate emotional responses, leading to changes in behavior, arousal, and their neural correlates. The link between arousal and loss aversion was demonstrated earlier in the project period.²⁴ On average, participants in this study were more aroused (as measured by skin conductance response) to losses relative to gains, and the difference in arousal to losses versus gains correlated with behavioral loss aversion. When participants were instructed to use a cognitive regulation strategy (i.e., “think like a trader”), behavioral loss aversion and physiological response were reduced. The amygdala is known to be involved in mediating the effect of emotional arousal in a variety of contexts, and this project serves as further evidence that behavioral loss aversion demonstrated in choices may indicate arousal related responses to outcomes.²⁵ No relationship was found between arousal and risk sensitivity.

²⁴ Sokol-Hessner, P., Hsu, M., Curley, N. G., Delgado, M. R., Camerer, C. F., & Phelps, E. A. (2009). Thinking like a trader selectively reduces individuals' loss aversion. *Proceedings of the National Academy of Sciences*, *106*, 5035-5040.

²⁵ Sokol-Hessner, P., Camerer, C. F., & Phelps, E. A. (2013). Emotion regulation reduces loss aversion and decreases amygdala responses to losses. *Social Cognitive and Affective Neuroscience*, *8*, 341-350.

A follow-up study to examine the causal relationship between arousal and loss aversion involved administering propranolol (i.e., a nonselective beta-adrenergic receptor antagonist (beta blocker)), which was used to target amygdala striatal circuitry and control loss aversion.²⁶ Propranolol has been shown to reduce the effect of emotion on memory by targeting the amygdala and diminishing its influence on the hippocampus. Baseline medical history, physical exam, and an electrocardiogram (EKG) were conducted with 47 participants, who were then randomly assigned placebo or 80 milligrams of propranolol²⁷ prior to a risky decision modeling task on two testing days with approximately 1 week in between visits. Findings suggest that the effectiveness of propranolol in reducing loss aversion is related to BMI: those with lower BMI showed a significant reduction in loss aversion. No relationship was found between modifying arousal and risk sensitivity. This study provides the first causal evidence that the amygdala-mediated arousal response (and/or its neural mediators) is not just selectively, but causally linked to loss aversion through beta-adrenergic pathways in the brain.²⁸

Acute Stress and Model-Free versus Model-Based Reinforcement Learning

Daw et al. described a dual-system for reinforcement learning.²⁹ The model-free system develops preferences for actions using temporal difference and relies on dopaminergic prediction error signals. The model-based system utilizes an internal model of environment to prospectively evaluate actions and is a prefrontal central-executive dependent system.

Phelps investigated how acute stress affects model-based and model-free reinforcement learning to demonstrate separability of these value systems using a two-step decision task designed by Daw et al.³⁰ The two-step task evaluated how bottom-stage feedback affects top-stage choices. The model-based index is equal to the size of the interaction of reward and transition for each individual. The model-free index is equal to the size of the individual's main effect of reward. Subjects in Phelps' study using a model-based strategy took the transitions into consideration, whereas subjects using a model-free strategy did not take transitions into account. Most subjects used a combination of strategies across the tasks. This experiment demonstrated an interaction between memory load and interaction on the last trial, meaning that model-based reinforcement learning requires working memory.³¹

Phelps examined the effect of acute stress on the reinforcement learning task to determine how previous reward and previous transition type determined current first-stage choice. A sample of 56 subjects was randomly assigned to a stress condition (i.e., cold pressor task) or a control. The protocol sequence included cortisol measure, operation span (test of working

²⁶ Sokol-Hessner et al. (submitted).

²⁷ This is the highest dose allowable for propranolol.

²⁸ Sokol-Hessner et al. (manuscript in preparation).

²⁹ Daw, N. D., Niv, Y., & Dayan, P. (2005). Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, *8*, 1704-1711.

³⁰ Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P., & Dolan, R. J. (2011). Model-based influences on humans' choices and striatal prediction errors. *Neuron*, *69*, 1204-1215.

³¹ Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A., & Daw, N. D. (2013). Working-memory capacity protects model-based learning from stress. *Proceedings of the National Academy of Sciences*, *110*, 20941-20946.

memory), cortisol measure, cold pressor test (or room temperature for control), cortisol measure, 200 trials of the two-step decision task, and a final cortisol measure. Results indicate a difference in cortisol between pre- and post-cold pressor task related to diminished model-based choices, but no effect was found on model-free choices. Those with higher working memory capacity were able to attenuate the effect of the acute stress.

In conclusion, these experiments demonstrated that acute stress response attenuates model-based contribution to learning, but not model-free contributions. This finding is consistent with the presumed dependence of the model-based system on prefrontal cortex-dependent executive resources³² and the deleterious impact of acute stress on working memory capacity and/or executive function.³³ Evidence from Phelps' and others' studies elucidate the relationship between stress and controlled prefrontal cortex dependent processing in decision making.

Self-Regulation As a Biological Mechanism for Excess Weight Gain in Low-Income Toddlers

Alison L. Miller, PhD, and Julie C. Lumeng, MD, University of Michigan, Ann Arbor

This longitudinal study is examining self-regulation of toddlers and parenting co-regulation. Data have been collected for phase I of the study on a sample of 180 low-income children who are 48 percent female, 51 percent White, 25 percent Black, 24 percent biracial, and 14 percent Hispanic/Latino. At age 21 months, 30 percent of the sample was overweight or obese based on Centers for Disease Control and Prevention (CDC) criteria. The design includes data collection at 21, 27, and 33 months of age. Physiological measures of the child and the parent include diurnal cortisol, alpha-amylase, and observational coding. Data collection is ongoing.

In an earlier study, Miller and Lumeng found that lower morning cortisol levels and a flatter pattern of cortisol decline over the course of the day were associated with being overweight at age 4 among low-income children. This finding prompted the current SOBC-funded study. The current study data suggest that child weight at 21 months predicts cortisol patterns over time. Being overweight at 21 months predicts a lower cortisol intercept and a flatter cortisol slope at 33 months. Additionally, being overweight at 21 months predicts a reduced cortisol intercept and a flatter slope from 27 to 33 months. At 33 months, the children who were overweight at 21 months have a significantly lower morning cortisol level and the level remains flat throughout the day, compared to a normative cortisol pattern for the children who were not overweight at 21 months. However, the data are complex and do show some bidirectional associations. A lower cortisol intercept and a flatter cortisol slope at 27 months are both associated with greater weight gain from 27 to 33 months. More analyses are needed of these preliminary data.

³² Otto et al., 2013.

³³ Schoofs, D., Preub, D., & Wolf, O. T. (2008). Psychosocial stress induces working memory impairments in an *n*-back paradigm. *Psychoneuroendocrinology*, 33, 643-653.

Preliminary data on the mother-child co-regulation show that mother and child morning cortisol and alpha amylase levels are positively associated at 21, 27, and 33 months. A mother's high morning cortisol level predicted a more rapid decline in child cortisol across the day at each age. It is possible that the mother-child physiologic synchrony is stronger in more sensitive dyads. Parenting behavior was also observed and coded. Mothers showing more positive affect during free play had children who showed less negative affect in the no-touch cookie task (i.e., they were better able to self-regulate). Mothers who were more sensitive and had a positive affect had children with higher morning cortisol and a more rapid decline throughout the day. Also, mothers who were more sensitive and had a positive affect had children who gained less weight between 21 and 33 months. The patterns are stronger in normal weight versus overweight toddlers.

Miller and Lumeng used a variety of measures to assess the developmental trajectory of food-related self-regulation. At 33 months, 63 percent of children failed the ability to delay gratification task and 71 percent failed the tongue task (i.e., hold a treat on your tongue but do not eat it). Some lasting individual differences over time were observed. Children at all ages tended to be more upset while waiting for a cookie than while waiting for a toy, although the percentage upset while waiting for a cookie decreased over the three data collection time points. Developmental changes are identifiable, and stable individual differences emerge early. These differences may vary across food and non-food self-regulation tasks.

In summary, preliminary data show a bidirectional association between physiological effects and weight gain. Being overweight at 21 months is associated with later atypical physiological regulation patterns (i.e., cortisol and alpha amylase), and physiological regulation is associated with later weight gain. Mother behavior is associated with child behavioral and physiological regulation and later weight gain. The findings differ, however, by child weight status. These early analyses indicate implications that should be considered in future research: food-related self-regulation tasks are developmentally sensitive, child factors contribute very early in the lifespan, and parenting is an important co-regulatory process that can be targeted and changed through intervention.

Poverty, Stress, and Discounting: A Mechanism for Behavior Change?

Johannes Haushofer, PhD, Princeton University

The working hypothesis of this project is that poverty affects psychological and neurobiological outcomes (i.e., stress), and, in turn, stress affects decision making. Two studies have been conducted on the link between poverty and stress, and two studies are being conducted on the link between stress and decision making.

Poverty and Stress

The first study demonstrated that negative income shocks in the form of rainfall had an effect on cortisol in farmers.³⁴ Rainfall in Kenya is random and affects the farmers' income, which affects the farmers' cortisol levels. This is not true for non-farmers.

The second study examined whether alleviating poverty can reduce stress. GiveDirectly is a nonprofit organization that provides unconditional cash transfers to poor households in Kenya. Haushofer and Shapiro conducted an RCT of 1,500 households that included treatment, spillover, and control villages. The treatment households received cash transfers of either \$400 or \$1,500. The investigators shared a pre-analysis plan publicly prior to the study and then conducted the analyses as planned.³⁵

Cortisol levels were significantly reduced among those who received the large \$1,500 cash transfer, but not for those receiving the small \$400 cash transfer. Patterns were similar for self-reported stress, although the measure was noisier.³⁶ In short, the RCT showed that poverty alleviation reduces self-reported stress and cortisol levels.

Differences were seen in other welfare indicators as well. There were large reductions in domestic violence among households receiving the cash transfers, regardless of whether the male or female was the recipient of the transfer. In addition, there was a spillover effect, that is, nonrecipient households in the same villages as recipient households showed lower levels of domestic violence.

Stress and Decision Making

The first of these laboratory experiments used hydrocortisone to manipulate subjects' cortisol levels and tests of temporal discounting to measure decision making. Results indicate that hydrocortisone administration increases temporal discounting. In other words, the subjects were more impatient when they had higher cortisol levels.

The second study examined whether there is a direct effect of poverty on decision making. Subjects participated in a laboratory experiment to assess temporal discounting after negative income shocks administered in the laboratory. Findings indicate that the participants discount more after a negative income shock, even when total wealth is kept constant.

The Busara Center for Behavioral Economics was established in Nairobi, Kenya, as a part of this work.³⁷ The Center is a public good, and researchers are invited to use it. The Center has a

³⁴ Chemin, M., de Laat, J., & Haushofer, J. (2013). Negative rainfall shocks increase levels of the stress hormone cortisol among poor farmers in Kenya. Social Science Research Network Working Paper.

³⁵ The pre-analysis plan is available at <https://www.socialsciregistry.org/trials/19>.

³⁶ Haushofer, J. & Shapiro, J. (2013). Household response to income changes: Evidence from an unconditional cash transfer program in Kenya. Available from http://www.princeton.edu/~joha/publications/Haushofer_Shapiro_UCT_2013.pdf.

³⁷ See <http://www.busaracenter.org>.

computer cluster with touch screen machines, which are more easily accessible for those who are illiterate, and a subject pool of 5,000 people who are recruited via text messages.

Summary

Key findings and their implications were presented in a recent publication in *Science*³⁸:

- Negative income shocks increase levels of cortisol and self-reported stress.
- Positive income shocks decrease levels of cortisol and self-reported stress.
- Pharmacological elevation of cortisol levels increases temporal discounting.³⁹
- Negative income shocks increase temporal discounting.⁴⁰
- These findings indicate that poverty may perpetuate itself by increasing stress and temporal discounting.

Stress and Stress Reactivity Measures and Targets

Continuous Measurement of Stress in the Mobile Environment via Wireless Physiological Sensors

Santosh Kumar, PhD, University of Memphis

Kumar presented ongoing work on the continuous measurement of stress and physiology that is supported by the NIH and the National Science Foundation.⁴¹ In 2007, the Genes, Environment and Health (GEH) program at the NIH called for field-deployable tools for measures of environmental exposures. Mobile tools are key for measuring environmental exposures because such exposures need to be measured frequently and in the natural context in which they occur. Multiple variables are interrelated and need to be measured simultaneously. Mobile sensors with similar properties are needed for measuring behaviors in the field.

Behavioral scientists want to study causal relationships among multiple variables (e.g., behaviors, contexts, exposures, stress states). Measuring behavior in the natural environment can be more challenging than pure physiological parameters, and therefore validation is difficult. However, innovations in hardware, software, data modeling, processing, and inferencing are enabling the development of platforms for continuous, unobtrusive measurement of behaviors in the natural environment. For example, social exposures can be measured with a microphone, visual exposures can be measured with glasses, geographic exposures can be measured using global positioning service (GPS) technology, and multiple

³⁸ Haushofer, J. & Fehr, E. (2014). On the psychology of poverty. *Science*, 344, 862-867.

³⁹ Cornelisse, S., van Ast, V. A., Haushofer, J., Seinstra, M. S., Kindt, M., & Joels, M. (2013). Time-dependent effect of hydrocortisone administration on intertemporal choice. Available from http://www.princeton.edu/~joha/publications/Cornelisse_et_al_2013.pdf.

⁴⁰ Haushofer, J., Schunk, D., & Fehr, E. (2013). Negative income shocks increase discount rates. Available from http://www.princeton.edu/~joha/publications/Haushofer_et_al_Negative_Income_Shocks_2013.pdf.

⁴¹ Plarre, K., et al. Continuous inference of psychological stress from sensory measurements collected in the natural environment. Paper presented at the ACM/IEEE International Conference on Information Processing in Sensor Networks (IPSN), Chicago, IL, April 12-14, 2011. (12 pages).

behaviors (e.g., exercise activity, smoking, eating) can be measured with a watch sensor and/or smartphone.

Kumar and colleagues⁴² have developed a suite of biosensors for measuring stress and behavior using AutoSense devices designed and built by Dr. Emre Ertin (Ohio State University). The suite of platforms includes chestband sensors, smartwatch sensors, and android smartphones. The devices upload data wirelessly in real time and use batteries that last 10 days. The first generation of these sensors was tested on 30 daily smokers and 42 drug users. Sensor data enabled the researchers to develop models of physiological stress and cocaine usage and preliminary models of smoking and conversation.

The sensors were also tested on 21 participants in a 2-hour laboratory-based validated stress protocol and in a 2-day field study of 10 to 14 hours. This study was conducted under supervision of Dr. Mustafa al'Absi at University of Minnesota Medical School. The goal of the study was to develop a stress model that can classify a given minute of data into a stressed versus nonstress class (i.e., in this particular minute, was the subject exhibiting a stress response or not). Electrocardiography (ECG) responses to laboratory stressors indicate higher levels of stress relative to baseline before, during, and after public speaking and mental arithmetic, and during the cold pressor task.

The model was then developed to be robust to noises in the natural environment. Even though heart rate variability has been used as a measure of stress, it is unclear whether it should be the only measure of stress. The team investigated several features and selected seven as the most informative, including features from respiration (e.g., inhalation duration, respiration duration, inspiratory/expiratory ratio, stretch, minute ventilation) and features from ECG (e.g., heart rate variability and respiratory sinus arrhythmia). To account for between-subject differences and to ensure the model was calibrated to the biases and reactivity for the individual, the features were normalized for each subject. Noise filtering techniques were used to account for background noise in the natural environment. Another modification for the field environment was that the no stress decisions are made during physical activity because currently there is no way to discern a stress response due to physical activity versus some other stressor. Initial tests of the field model demonstrated that the sensor classified stress versus nonstress responses with 93 percent accuracy. At this point in the project, the sensor was still evolving and the housing for the sensor needed improvements.

Because self-reported measures of stress tend to last longer than what physiological measures indicate (e.g., heart rate variability recovers quickly from cold pressor test, but a person may report feeling stressed much longer), Kumar and colleagues developed a binary hidden Markov model to estimate the stress state in the brain using physiological variables. Parameters for each individual are estimated and then used to predict how a person will respond when asked

⁴² Ertin, E., et. al., AutoSense; Unobtrusively wearable sensor suite for inference of onset, causality, and consequences of stress in the field. Paper presented at the ACM International Conference on Embedded Networked Sensor Systems (SenSys), Seattle, WA, Nov. 1-4, 2011. (14 pages).

to self-rate stress. Stress sessions in the laboratory were used to predict each of 14 self-reports and yielded a median correlation of 0.72. In the field, the model was used to predict the average stress self-rating of each subject over 2 days, and the correlation coefficient was 0.71. These can be compared to correlation coefficients of self-report and measures of cortisol level of 0.3-0.4.

Preliminary data from a current field study of 30 smokers and 42 drug users indicate high levels of stress during stop-and-go driving in traffic and prior to self-reports of smoking, drug use, or craving. The densest levels of stress are observed during transit.

The long-term goal of these technologies is to employ the models to detect stress that predicts a near-term engagement in a harmful behavior and then deploy a just-in-time adaptive intervention

STRAIN: An Automated System for Assessing Lifetime Stress Exposure

George Slavich, PhD, University of California, Los Angeles

The goal of Slavich's research is to understand the full set of psychological and biological mechanisms that link stress to poor health and to develop tools that help investigators conduct similar research in a high-quality, cost-effective manner. Americans suffer from more chronic disease burden than their counterparts in 15 other economically comparable developed countries.⁴³ What is most troubling is that the majority of the U.S. health burden is attributable to preventable conditions such as infant mortality, obesity, and heart disease. Traditional health behaviors, such as physical inactivity, alcohol consumption, and smoking, are well known to contribute to chronic disease. However, the leading risk factor for chronic disease-related mortality is actually high levels of social stress.

Gold Standard Measure of Life Stress

One system for assessing stress is the Life Events and Difficulties Schedule (LEDS). The LEDS system involves a 2-hour semi-structured interview that gathers information about different kinds of acute life events and chronic difficulties occurring in 10 life domains, including health, housing, finances, work, education, and relationships. An independent panel of expert raters reviews the interview material to form a consensus judgment regarding each stressor's degree of contextual threat, taking into account the unique features of each individual. Raters use a 500-page manual with more than 5,000 pre-rated case vignettes and extensive rules and criteria for rating life stress, distinguishing between acute and chronic stress.

There are several advantages and disadvantages of the LEDS system. Systems such as LEDS provide investigators with high-resolution stress data, but only over a very short period of about 1 to 2 years. Also, this assessment strategy is not consistent with current literature,

⁴³ National Research Council and Institute of Medicine. (2013). *U.S. Health in International Perspective: Shorter Lives, Poorer Health*. Washington, DC: The National Academies Press.

which supports the notion that health trajectories are influenced by stressors occurring over the entire life course.

LEDS Advantages

- Yields very rich stress exposure information, including contextual details about all recent stressors experienced and biographical details of the person
- Produces standardized, contextual threat ratings
- Differentiates acute versus chronic stressors
- Identifies stressor characteristics, including life domain and core social-psychological characteristics
- Assesses duration and timing of stress exposure in relation to different biological or health outcomes of interest
- Precisely identifies what happened (i.e., low “intra-category variability”)

LEDS Disadvantages

- Requires a lot of personnel (e.g., interviewers, raters, data managers) and extensive training
- Incurs huge incremental cost (i.e., about 6 hours and \$1,500 per additional participant)
- Has poor scalability
- Employs decentralized data storage
- Literature based on relatively few outcomes (primarily schizophrenia and depression)

Slavich and colleagues have hypothesized that stress occurring over the life course can become embedded in the body in different ways to affect health. For example, stress can shape social-behavioral repertoires, affect available coping resources, reinforce negative cognitive schemas, remodel anatomical and functional brain connectivity, alter the threshold at which the hypothalamic-pituitary-adrenal (HPA) axis and sympathetic nervous system (SNS) are activated, alter molecular signaling pathways, and alter the expression of hundreds of genes that influence all aspects of cognition, emotion, and behavior.^{44 45} To address the challenges of measuring stress exposures over the life course, Slavich and colleagues have developed the Stress and Adversity Inventory (STRAIN), which is the first automated system for obtaining a cost-effective, high-resolution measure of lifetime exposure to stress.

STRAIN: Stress and Adversity Inventory

STRAIN is an online assessment system that measures individuals’ lifetime exposure to different types of acute and chronic stressors that affect health.⁴⁶ STRAIN is not intended to be interchangeable with highly intensive, interview-based systems such as LEDS, but rather seeks

⁴⁴ Slavich, G. M. & Cole, S. W. (2013). The emerging field of human social genomics. *Clinical Psychological Science*, 1, 331-348.

⁴⁵ Slavich, G. M. & Irwin, M. R. (2014). From stress to inflammation and major depressive disorder: A social signal transduction theory of depression. *Psychological Bulletin*, 140, 774-815.

⁴⁶ More information about STRAIN, including literature and the technical manual, is available at <http://www.uclastresslab.org>.

to combine the reliability and sophistication of an interview-based measure of stress with the feasibility of a self-report instrument.

STRAIN assesses peoples' lifetime exposure to 96 different types of acute life events and chronic difficulties. Items are written in eighth-grade English and are presented one at a time on a computer screen. Additional options include having the questions read aloud and indicating responses with a touch screen via iPad. It can be self-administered or facilitated by an interviewer. For each stressor that a participant indicates experiencing, users are asked a series of tailored follow-up questions that ascertain the severity, frequency, timing, and duration of the stressor. The assessment yields 445 raw variables and 115 stress exposure summary scores and life charts, and it takes participants 25 to 35 minutes to complete. The stress exposure indices are created using assigned attributes to each core question. Current indices include exposure timing, stressor type, life domain, and social-psychological characteristics.

Slavich currently has STRAIN data on 10,085 participants from 32 studies spanning several populations, age groups, and in relation to all major outcome levels of analysis including neural, social, cognitive, emotional, physiologic, endocrine, molecular, and genomic processes. These data have been collected using the full and complete version of STRAIN. Slavich is currently validating a condensed version of STRAIN that aims to capture the most critical variables in 10 minutes or less.

Slavich has also worked on developing clear and compelling ways to graphically display data. For example, a contour plot of age and stress exposure severity (scale of 1 to 5) quickly displays the ages during the life course at which severe stress exposures occur. Another example is the use of a stress exposure life chart of age and density including the full sample, and then superimposing various sub-populations such as four randomly selected subjects or caregivers versus non-caregivers. Heat plots of frequency of stressors and a health outcome or grouping factor of interest (such as telomere length, interleukin-6 (IL-6) levels, tumor necrosis factor-alpha (TNF- α) levels, or BMI) are also visually instructive methods for displaying the stress data.

Predictive Utility Examples

Measures of lifetime stress exposure from STRAIN have predictive utility for outcomes in a variety of domains and levels of analysis. For example, a sample of 148 young adults (mean age = 19.3 years) was administered the STRAIN, Kessler 6 (a 6-item global measure of mental health), and Physical Health Questionnaire (a 14-item measure that assesses somatic and physical health complaints). Results of the analysis indicated that count and cumulative severity of stressors occurring across the young adult life course were both strongly associated with poorer physical and mental health, with some differential effect by stressor type.

A different sample of 112 young adults (mean age = 19.9) free of psychiatric or medical illness was randomized to control (N = 55) and stressor conditions (N = 57). Participants were administered STRAIN, Wisconsin Card Sorting Test (executive function), Beck Depression Inventory (depression), Patient Health Questionnaire (physical health), Kessler-6 (mental health), and pre-stressor anxiety measures of affective and self-reported perceptions of threat

to the upcoming stress condition. Analyses of the data indicated significant predictive associations between lifetime stress exposure as measured by STRAIN and the outcomes of physical health ($p < .001$), global mental health ($p < .001$), depressive symptoms ($p < .001$), executive function ($p < .026$), and anticipatory anxiety ($p < .034$).

Another study on lifetime stress and cancer-related fatigue involved a sample of 50 breast cancer survivors (mean age = 58), 30 of whom were experiencing persistent fatigue. Participants were administered the STRAIN assessment and the Short Form-36 vitality scale, a reliable and valid measure of energy and fatigue in the general population. Breast cancer survivors with persistent fatigue reported significantly higher levels of cumulative lifetime stress exposure, including more stressful experiences in childhood and adulthood, compared to the control group of nonfatigued cancer survivors. In an additional analysis, investigators computed a measure of stress burden, which took into account both a person's amount and duration of lifetime stress exposures. This analysis indicated that for every one-point increase in the stress burden measure, there was a 17.29 percent increase in the probability that the participant was experiencing persistent fatigue. These findings point to a novel risk factor for fatigue in the growing population of cancer survivors and suggest targets for treatment.

There is also predictive utility of STRAIN measures for biological responses to stress. A study of 46 healthy older adults (mean age = 63), of whom 22 were dementia caregivers, included the STRAIN assessment and a measure of IL-6, which is a risk factor for coronary heart disease. Greater cumulative stress severity predicted higher levels of circulating IL-6 in this sample. The same sample was used in a study of lifetime stress exposure and biological aging, as measured by leukocyte telomere length. Results indicate that telomere length decreased from time 1 to time 2 for participants with high cumulative severity of stressors, and that these effects were robust while controlling for age, BMI, and caregiver group status.

There is moderate correlation between the predictive utility of STRAIN indices and other measures of stress including the Lawton Appraisal (impact of caregiving), Wheaton checklist (social stress), the Perceived Stress Scale, and the NIH Life Chart. However, STRAIN appears to be a better predictor of many of the biological responses of interest, particularly IL-6 and changes in telomere length.

Moderating Factors

Slavich and colleagues have just begun examining moderating factors. In a sample of 71 high- and low-stress caregivers (mean age = 62) who were free of diabetes and metabolic syndrome, analyses indicated that both greater lifetime stress burden and pessimism independently predicted higher metabolic risk. However, these main effects were qualified by a significant interaction effect of lifetime stress exposure and pessimism, whereby pessimism significantly exacerbated the negative effects of lifetime stress exposure on metabolic risk score, after adjusting for age, income, and caregiver status. These findings suggest that pessimistic beliefs are a possible treatment target for reducing stress-related disease burden.

Another study of a previously mentioned sample (N = 148 young adults) included the Heartland Forgiveness Scale, which is a measure of the emotion-focused coping strategy of forgiveness. This analysis found that forgivingness moderated the effects of lifetime stress exposure on mental health in a graded fashion, whereby lifetime stress exposure was most strongly associated with poor mental health in persons exhibiting low forgivingness.

A final example of moderating factors involved a sample of 31 healthy young adult females (mean age = 19) who were administered the STRAIN assessment and the Social Provisions Scale (a 24-item self-report scale that measures the extent to which social relationships are perceived as supportive), as well as an fMRI session in which they believed they were being negatively evaluated by another participant. Two subscales included attachment (i.e., assessing emotional closeness) and reliable alliance (i.e., perception of relationships as being able to provide concrete assistance). Analyses indicate that more lifetime stress exposure was associated with greater neural responses to negative versus neutral social feedback and that these associations were significantly strong in persons reporting low social attachment.

Question and Answer

Nielsen inquired about the qualities of STRAIN that give it good predictive utility, particularly when compared to other measures of stress. Slavich reported that STRAIN is very structured in terms of the questions it asks. The questions have a severity threshold built into them to get people to focus on acute life events and chronic difficulties that are moderately or strongly impactful. The text of STRAIN is written to minimize confusion about what the questions are asking, possibly more so than in other measures.

Compared to the cost of the LEDS administration—approximately \$1,600 per additional participant—STRAIN is incredibly cost-effective. STRAIN is a computer-based model run on secure servers. The per-person cost can be as low as a few dollars, depending on the size of the study. Slavich clarified that this is “STRAIN v1.0” and that a long-term goal is to include additional follow-up questions that a LEDS interviewer might ask to produce a separate set of stressor severity ratings that are based on the objective qualities of the stressor (i.e., as opposed to participants’ self-reported severity ratings). The LEDS system includes extensive information about how to rate the severity of different types of acute life events and chronic difficulties, and Slavich hopes to build this intelligence into the STRAIN system going forward.

Richard Heyman asked about the advantages of objective measures of stress versus self-report. Slavich shared that he has a long history of experience with LEDS and similar systems, and he feels strongly about the importance of objective measures of stress. However, he mentioned that all effects of stress on health are “cognitively mediated.” In other words, when it comes to downstream consequences, the health effects of all stressors are shaped by an individual’s experience of the stressor. STRAIN attempts to identify the presence of stressors that are most impactful. At the same time, there needs to be a balance between assessing the objective nature and impact of the stressors that people experience while also assessing the unique perceptions that individuals have of those stressors.

Slavich noted that there are a host of issues that could be discussed in terms of the validity of retrospective data. The correlation of test-retest reliability for STRAIN is 0.80. It appears that people are mostly recalling the same types of experiences at baseline and 6 months later, indicating that it is a valid snapshot and perhaps more valid than a short checklist measure of stress. The one-at-a-time question presentation on the screen facilitates the test-retest reliability of STRAIN in that people do not realize they are completing such a long and detailed interview (if printed out on paper, STRAIN would be 185 pages long).

Stress and Stress Reactivity Group Discussion

Facilitators: Susan Czajkowski, PhD, National Heart, Lung, and Blood Institute; Wendy Weber, PhD, National Center for Complementary and Alternative Medicine; Elizabeth Phelps, PhD, New York University

Multiple Perspectives of Stress: Talking Across the Aisle

Czajkowski opened the discussion session by asking the group to focus on how to reconcile the multiple conceptualizations of stress and stress reactivity in the context of behavior change. Stress can be a stimulus, dependent variable, or moderator. Stress exposures can be chronic or acute and can be measured objectively versus self-report over different time frames. Stress decay and recovery are important related constructs. Czajkowski asked for feedback from the grantees and invited speakers on the necessary next steps to move the field forward, reconcile these various conceptualizations and perspectives, and identify when and how a stress mechanism becomes a target for behavioral intervention.

Phelps commented that her own work on manipulating and testing acute stress in a laboratory setting has demonstrated its importance. For her presidential symposium at the Association for Psychological Science's (APS) 2014 Annual Convention, Phelps compiled a diverse panel of speakers to address the roots of and interventions for stress from neurobiological, cognitive, health, and developmental perspectives.⁴⁷ One thing that struck her at the symposium was how the different fields studying stress had not previously communicated and did not know each other. The social psychology world of stress is not talking to the neurobiological world of stress. The two invited speakers on stress at today's meeting—Kumar and Slavich—defined stress differently and viewed it through different lenses. Many investigators are interested in addressing these conceptual issues, but the first step has to be interdisciplinary communication and collaboration among those working on similar constructs.

Nielsen commented further about the different perspectives presented by Kumar and Slavich. In her opinion, Kumar is not working in a stress framework per se; rather he is focusing on the moments leading to health behavior failure. Kumar's sensors measure the reactivity to stress that is the target of intervention. The STRAIN assessment is focused on lifetime stress exposures and not as much on reactivity. Both stress exposures and reactivity are important. What needs to be clear in any conversation is what is being targeted for intervention—some

⁴⁷ See more information at <http://www.psychologicalscience.org/index.php/video/taking-it-easy-isnt-easy.html>.

capacity within a person to react to stress or a change in an environmental exposure? Either could be appropriate intervention targets depending on the behavior and context.

Subjective and Objective

Levenson drew a distinction between the simplicity of defining stress in animal models—you shock a rat and the stress response is heart rate acceleration, squealing, etc.—and the complexity of defining stress in humans. Acute stress imposed on humans in a laboratory setting can induce myriad responses: laughter, fear, sadness, etc. Defining stress in humans as the acute stimulus is not adequate. The definition of stress as experienced by humans is inextricably linked to how they think and feel about it. The environmental stimulus and biological response in humans still need to be characterized, but all of it will be moderated by the set of cognitive and emotional responses employed by individuals in reaction to the stress. Both physiology and cognitive and emotional moderators of stress response are important.

Slavich asserted that the theories and concepts of stress and refinement of stress are only as good as the measurement strategies employed. Historically, stress has been measured crudely, and there has not been sufficient refinement of the construct. It seems as though pieces of the concept of stress have been defined arbitrarily in conjunction with psychologically imposed ideas of what stress is. Slavich's approach, instead, is to follow the biological systems' response to stress. The question becomes "what does the immune system care about?" For example, what is the impact of interpersonal loss, social isolation, and hostility in the environment from the perspective of the immune system? This biologically informed way of thinking about how to measure stress is very different from categorizing aspects of stress according to what psychologists think the categories should be.

Insights from NCI Meeting on Negative Affect and Stress

Ochsner recounted his experiences at a workshop organized 2 years earlier by NCI on the relationship between negative affect and stress. From his perspective, it seemed that participants concluded that "stress" is the term used to mean the stimulus of the negative affect. Nielsen and Klein, who also attended, noted differences between the fields of negative affect and stress that were apparent at the meeting. The stress literature appears to be tied to the long-term cumulative impact on health outcomes. Another related concept—threat sensitivity—concerns the propensity to attend to negative stimuli. There is a link between negative affect and stress, but they are measured and studied differently. One point of dissension that Klein observed at the NCI meeting was that affect can be characterized by valence but stress typically is not. Klein noted that affect can be positive or negative and that there are researchers studying both aspects. Yet stress has not typically been thought of positively. Recent research on resilience is taking the stress literature in a different direction.

Producers and Consumers: Lock the Door?

Haushofer agreed that there is a great deal of conceptual work to be done and wondered about the utility of convening a mix of people whose daily work it is to define and measure stress (producers) and those who consume these definitions and measures for use in ongoing studies that involve stress (such as his own). As a consumer, Haushofer believed he would benefit from

understanding directly from those working on stress how it is currently being defined and measured. King observed that convening such a group would be useful to the producers as well because hearing from consumers could help them refine concepts and measures and focus on aspects of stress that are useful for behavioral interventions, such as those focusing on endocrine aspects of stress or perceptions of stress. There does not appear to be consensus on what concepts and measures are important for which type of person or intervention. Weber suggested that such a group could define the minimum data set required for studying stress and then identify some common measures.

Grantee Updates: Session 3

Using Media to Explore Mechanisms of Behavior Change among College Students

Megan Moreno, MD, MPH, Seattle Children's Research Institute and University of Washington

College students are generally healthy but the morbidity and mortality that does occur in this population tends to be linked to substance and alcohol use. Young adults in college tend not to seek health care and when they do, they are not typically screened for substance and alcohol abuse. This study is examining the utility of social media to identify young adults at risk for substance and alcohol abuse and also as a tool for changing behavior.

The study sample includes 330 college students from two universities. Project staff complete monthly Facebook profile evaluations of each participant and use a combination of scheduled interviews and interviews prompted by new posting behavior on Facebook. Facebook use is still very high among college students: 98.4 percent of sophomores and 98.8 percent of juniors maintain Facebook profiles. Other social media outlets, such as Twitter, LinkedIn, YouTube, and Instagram, are used to a lesser degree.

In this sample, displayed substance use references are common on college students' Facebook profiles and increase with each year of college. Data were collected at three time points: before starting college, year 1, and year 2. Data collection for year 3 is not complete. Facebook profiles with alcohol use references jumped from 20 percent before college to 68 percent in year 2. Moreno expects further increases in year 3, especially as students turn 21. Marijuana references increased from 3.2 percent before college to 12.0 percent in year 2.

Facebook displays of substance and alcohol use have meaning, but the type of display matters. For example a "like" on a Facebook page is not very meaningful compared to status updates, photos, and cover or profile photos. The positive predictive value of a photo referencing alcohol use for actual alcohol use in the past 28 days is 100 percent if the photo is used as the cover or profile picture. The positive predictive value for four or more binge episodes in the past 28 days for a general photo is 47.4 percent and 100.0 percent if the photo is used for the cover or profile. These are important distinctions to consider when using big data techniques—these

fine-grained differences between likes, status updates, and where a photo is posted and how they are differentially predictive of behavior might be missed.

Facebook is perceived by college students as being influential, but it is complicated. Moreno developed a Facebook influence model using concept mapping with input from college student focus groups. Thirteen different clusters in four domains emerged as the elements that make Facebook influential. The domains and clusters include connection (connection to people, far reaching, fast communication, business and promotion, accessible and adaptable, data and information); identification (identity expression, influence on identity); comparison (curiosity about others, Facebook establishing social norms); and Facebook as an experience (distractions, positive experiences, negative experiences). There are areas where the clusters overlap between what they describe and theories of influence, such as social norms, social information processing theory, media practice model, and media richness theory.

In summary, Facebook remains popular with the older adolescent population and is a common place for college students to display references to substance use. These references have validity to the profile owners, and the context of the reference is important when discerning its value in predicting behavior. These references may influence others who view them via several mechanisms of behavior change, as described in the influence model.

Determinants of Adolescent Exercise Behavior: Toward Evidence-based Intervention

Meike Bartels, PhD, VU University Amsterdam

This study involves an international collaboration between Netherlands Twin Registry (NTR) and the Vermont Family Based Approach at the University of Vermont with Principal Investigator James Hudziak. The project comprises four studies: (1) longitudinal survey of twins and their family members between ages 7 and 100; (2) laboratory study of 227 complete adolescent twin pairs and 38 of their singleton siblings (age: 17.2 ± 1.2 , range 12-25); (3) telephone interview and actigraphy for monozygotic twins who were found to be discordant for physical activity ($n = 100$ pairs); (4) and an exercise intervention for families in the Vermont program. The major achievement of the past year is extensive data collection and the continuation of the longitudinal study.

The primary preliminary finding based on cross-sectional survey data of 21,000 twins is that no consistent association between regular leisure time exercise behavior and BMI was observed from age 7 to age 18. This finding implies that the relative importance of a lack of regular exercise among youth as a major cause of the obesity epidemic should be reconsidered. The longitudinal data were also examined to determine causality.

A laboratory protocol was conducted with 227 adolescent twin pairs and 38 of their singleton siblings. Data collection concluded a few months ago. Variance in physical ability was explained by a combination of genetic factors and person-specific environmental factors. Genetic factors

are important in explaining part of the variance in physical activity, especially for oxygen uptake, which is one of the most important aspects in all physical ability, and performance on tasks of flexibility and hand grip strength.

An archival study of a clinical population was used to inform the development of an exercise intervention for child and adolescent psychiatric patients. Child MET scores and withdrawn/depressed scores significantly differed across exercise categories (i.e., no exercise, individual-only, team only, and individual and team sports). Child MET scores and team sport participation are each inversely related to withdrawn/depressed scores. Electronic sedentary behavior (e.g., TV, computer gaming, internet use) is positively associated with withdrawn/depressed scores.

The team developed an incentive-based electronic exercise intervention for adolescent psychiatric outpatients aged 13 to 17. Participants wear a physical activity monitor and receive daily text reminders to log exercise behavior, perceived exertion, daily hassles, and mood via an online survey. They are eligible to earn monetary rewards for participation and behavior change, respectively. The final study involves a telephone interview and use of an actigraph for monozygotic twins selected because they were extremely discordant for physical activity. These last two studies in the project are ongoing.

In summary, preliminary findings and lessons include the following:

- Voluntary exercise is increasingly influenced by heritability with age.
- Physical strength and physical ability are moderately heritable.
- There is a significant association between withdrawn/depression scores and voluntary exercise behavior in a clinical sample.
- Surprisingly, there is no association between voluntary exercise behavior and BMI.
- A large-scale actigraphy project is feasible in a genetically informative design.

Enhancing Adherence Behaviors in Cancer Patients During Radiation

Eileen Shinn, PhD, University of Texas MD Anderson Cancer Center

Shinn's parent grant is focused on adherence to exercise under conditions of extreme stress. The incidence of pharyngeal cancer has doubled in the past 20 years because of the human papillomavirus (HPV) epidemic. The 5-year survival rate for diagnosis at stage III or IV is 61-78 percent.

Typical treatment consists of high-dose radiation of the head and neck including key swallowing muscles. The radiation scars connective tissue and blood vessel linings, resulting in 22-39 percent of patients not being able to swallow normally after treatment. Swallowing exercises during radiation treatment have been shown to be effective in preventing dysphagia. However, 55 percent of patients do not adhere to the exercise regimen. The exercises are painful, and it is difficult for the patients to see the long-term benefits when faced with more immediate

concerns at the time of treatment (e.g., survival, pain, fatigue). One-half of the patients will not even try the exercise one time, and One-half end up requiring a feeding tube.

The parent study has enrolled 398 patients in the study out of 518 approached and randomized them to treatment and control. Data were collected at baseline and at 6-week follow-up, and is still being collected at 1-year follow up. Preliminary data demonstrate a statistically significant difference between treatment and control patients in the average number of sets of exercises performed during radiation. More importantly, all swallowing outcomes are improved in patients randomized to the treatment group, which holds true even when controlling for stage of disease, level of radiation and chemotherapy, and location of tumor.

In considering what behavior mechanisms might be fueling the behavior change, the team conducted semi-structured phone interviews with 27 of the treatment patients who had changed from nonadherent to adherent during the study, which yielded two main factors: feeling of accountability to a friendly counselor and knowing they would be monitored by the counselor. The SOBC administrative supplement is testing these two mechanisms for behavior change in a study in which patients were randomized to three groups and asked to perform a set of sham exercises. The groups included high- versus low-touch accountability and high- versus low-touch monitoring. In the experimental condition, high accountability was conceptualized using Tetlock's accountability paradigm. Patients were told, "We are being considered as one of the testing sites for a major study and being chosen would be a great honor for the Head and Neck Center team at MD Anderson. One of the ways we will be evaluated for the trial network is if we are able to show that our patients have high adherence to the new exercises. It is very important to us that you try your hardest to do as many repetitions as you can."

To date 92 of the 120 patients who consented to participate have completed the supplement study protocol. Preliminary data indicate that high accountability and human monitoring resulted in increased adherence ($p = .02$). The effect of the experimental condition was not influenced by side effects or number of radiation sessions.

If it can be shown through this supplement study that the low-touch accountability condition is sufficient for behavior change, then the importance of doing the exercises could be communicated to patients through a digital platform (e.g., video chat) rather than in person. This could have implications for sensor research, clinical decision making, and future health policies. Patients could use wearable sensors that communicate with clinicians to track adherence and check in with their doctors face-to-face on a less frequent basis.

The study team is currently participating with the Cyberinfrastructure for Comparative Effectiveness Research (CYCORE) project, a cyber infrastructure network that integrates patient collected data with MD Anderson's electronic medical records and CYCORE research

databases.⁴⁸ The data are then processed and analyzed securely to generate information that can inform clinical decision making.

Interpersonal and Social Processes Measures and Targets

Interpersonal and Social Processes: Measures and Targets

Robert W. Levenson, PhD, University of California, Berkeley

Levenson's work has primarily targeted dyads of married couples. Marriage is still something that almost everybody does at least once in their lives, but about one-half of marriages end in divorce. Levenson has also targeted dating couples, couples from different cultures, couples at different stages in the life course, caregiver-dementia patient couples, and lesbian and gay couples.

Levenson's work focuses on emotion and how it is used in a way that is relevant to behavior change. It is important to break down emotion in terms of basic processes that the emotional system has to take care of. Several aspects of emotion can be measured: emotional reactivity (magnitude, duration, latency), emotion regulation (up and down regulation, ability and typical practice), and emotion recognition (in self and in others and in terms of its sequelae such as empathy, compassion, pro-social behavior, manipulation, or exploitation). Studies of emotion regulation have typically focused on down-regulation. However, up-regulation of emotion is receiving more attention and can be important in couples because emotional signals need to be clear for the spouse.

Up to 98 percent of emotion regulation episodes take place in social contexts.⁴⁹ However, there is a clear bias toward intra-individual definitions and measures in existing research of emotion regulation. Less than 12 percent of studies published between 2001 and 2011 assessed emotion regulation in the presence of another person, and many of these were in the presence of an "imagined other."⁵⁰ Most of the emotion regulation research that focuses on dyads is in the mother-infant literature, particularly on behavioral measures of attunement and synchrony. There is currently no off-the-shelf measure of emotion regulation in dyads. Self-report measures have limitations, and they do not fare well when compared with objective performance-based measures. Self-reports appear not to be well calibrated in terms of how "big" a person's emotions are, how well the person is actually regulating, and how good he or she is at reading others' emotions.

⁴⁸ More information on CYCORE can be found at

http://cwphs.ucsd.edu/index.php?option=com_content&view=article&id=76%3Acycore&catid=34&Itemid=58.

⁴⁹ Gross, J. J., Richards, J. M., and John, O. P. (2006). Emotion regulation in everyday life. In D.K. Snyder, J.A. Simpson, & J.N. Hughes (Eds.) (pp. 13-35). *Emotion regulation in families: Pathways to dysfunction and health*. Washington, DC: American Psychological Association.

⁵⁰ Campos, J. J., Walle, E., Dahl, A., & Main, A. (2011). Reconceptualizing emotion regulation. *Emotion Review*, 3, 26-35.

Levenson's approach to measuring emotion in dyads is to maximize naturalistic aspects of interactions between couples while maintaining the rigors of a laboratory study. Dyads engage in 15-minute unrehearsed conversations in the laboratory about events of the day, a problem area, and a pleasurable activity. Continuous measures during the conversation include measures of physiology (autonomic and somatic nervous systems), behavior (coding of videotape), and subjective experience (participants rate their own and their partner's emotional state while watching the videotape of their conversation).

Levenson played a 2-minute excerpt of a videotape from a dyad conversation to illustrate some of the measurement issues. Emotional reactivity was measured in terms of coding behavior. One couple had 55 anger events in the full 15-minute conversation (3.7 anger events per minute, which is the fourth highest rate in the laboratory's history). Sequential probabilities can also be calculated—that is, the probability that the husband's anger is followed by the wife's anger or the wife's anger is followed by the husband's humor. These behavioral probabilities can give a sense of how couples exchange emotion. The patterns of exchange are predictive of the course of the marriage, health of the partners, and well-being of children in the family.

Physiological responses (e.g., heart rate, sweat gland activity, temperature) of each partner are measured continuously throughout the conversation. A composite of all measures can be normalized and averaged second-to-second to indicate the overall physiological cost of the interaction. Couples can become locked into a positively or negatively attuned state. Emotion regulation can be viewed in terms of time needed to cool down, which turns out to be very valuable and predictive of how the marriage will fare over time.

This work found evidence of significant long-term consequences of poor emotional reactivity, regulation, and recognition in married couples. High levels of anger, but not fear or sadness, predicted emergence of cardiovascular symptoms over decades of marriage. These results did not differ as a function of sex or age. The effects were specific as to emotion and type of symptoms and only emerged over time. The effect size was large. The odds ratio for anger and cardiovascular symptoms was 1.5, which is comparable to some of the odds ratios associated with abdominal obesity or lack of physical activity in predicting myocardial infarction. The association between anger behavior and cardiovascular symptoms was similar for husbands and wives and across middle-aged and older adults. Results held even when controlling for sociodemographic characteristics, health behaviors and problems, and marital satisfaction, all of which can have independent effects on cardiovascular health. These findings imply that angry marriages may present a significant hazard to cardiovascular health and point to a possible behavior intervention target.⁵¹

Levenson's ongoing research and these findings in particular have implications for future work on interpersonal and social process targets and measures for behavioral change:

- Interpersonal and social measures and targets and measures are potentially of great utility for the science of behavior change.

⁵¹ Haase et al. (under review).

- Emotions in dyads offer promising targets for behavior change efforts and can serve as useful measures of the impact of behavior change effects.
- Adopting these approaches will mean moving beyond the historical default of individual targets and individual self-report measures and adopting more interpersonal targets and measures.

Assessment of Interpersonal Mechanisms

Richard E. Heyman, PhD, New York University

Heyman and colleagues recently referenced Kurt Lewin's well-known assertion that there is "nothing so practical as a good theory" by adapting the notion to say that there is "nothing so practical as a good theory-testing tool."⁵² ⁵³ Heyman and colleagues have examined the assessment of interpersonal mechanisms in several contexts: couples, families and oral health, new parent intervention, and doctor-patient relationships.

Couples Observation

A 2001 review of more than 200 couples observation studies yielded several converging findings across interventions and methods of observation.⁵⁴ Distressed partners, compared with nondistressed partners, are observed to be more hostile; start conversations with hostility and maintain the hostility during the course of the conversation; are more likely to reciprocate and escalate hostility; edit less, resulting in longer negative reciprocity loops; are less positive; suffer more ill health effects; and are more likely to become withdrawn. Not a single study in this collection defined and measured constructs in the same way, which makes it very difficult to replicate findings and build a body of literature. About 10-15 minutes of observation time is needed to reliably code an interaction.⁵⁵

Based on these findings, Heyman provided recommendations to improve the assessment of interpersonal processes: (1) construct validity should be a prime concern because of the idiosyncratic combinations of micro codes and the context-dependent nature of validity; (2) reliability and agreement should be evaluated at the level of analysis; (3) multi-method assessments of constructs should be used; and (4) there should be experimental control of an analogue situation. Heyman provided examples from a study on emotion regulation in couples that involved bringing a couple into the laboratory and observing them while they discuss a high conflict topic to observe instances of anger in a fairly natural context. The couples completed ratings of experienced emotion while watching a video of themselves.

⁵² Lewin, K. (1951). *Field Theory in Social Science: Selected Theoretical Papers*. New York: Harper & Brothers Publishers.

⁵³ Heyman, R. E., Lorber, M. F., Eddy, M., & West, T. V. (2014). Behavioral observation and coding. In H. T. Reis & C. M. Judd (Eds.) (pp. 345-372). *Handbook of Research Methods in Social and Personality Psychology* (2nd Ed.). New York: Cambridge University Press.

⁵⁴ Heyman, R. E. (2001). Observation of couple conflicts: Clinical assessment applications, stubborn truths, and shaky foundations. *Psychological Assessment*, *1*, 5-35.

⁵⁵ Heyman, R. E., Chaudhry, B. R., Treboux, D., Crowell, J., Lord, C., Vivian, D., & Waters, E. B. (2001). How much observational data is enough? An empirical test using marital interaction coding. *Behavior Therapy*, *32*, 107-123.

Family Dysfunction and Oral Health

Heyman's work on family and oral health is based on the hypothesis that family dysfunction is related to oral health through two pathways: (1) dysfunctional family behavior leads to poor oral health behaviors, which leads to poorer oral health outcomes and (2) noxious family environments lead to negative oral health outcomes through the immune and endocrine systems. Noxious family environments can include characteristics such as partner aggression and abuse (i.e., physical, emotional, or sexual), couple conflict, child exposure to couple conflict and/or aggression, child maltreatment, and/or dysfunctional parenting. Results from Heyman's work indicate that a noxious family environment as observed by hostility and emotional and physical aggression is associated with poorer oral health for the adults and children.⁵⁶

Parent Intervention

Another study involved an intervention for parents to prevent noxious family environments. The intervention, *Couple Care for Parents*, can be delivered flexibly via video, workbook, and coaching and has been demonstrated in Australia to have excellent results.⁵⁷ Heyman and colleagues recently completed a RCT with American couples. The work is being extended with funding from the National Institute of Dental and Craniofacial Research (NIDCR) to include oral health outcomes and will measure and assess the mechanisms of behavior change. Repeated measures include observations of parent and child during a feeding task and partner interactions during a conflict task. Inter-parental conflict, diet, and brushing are moderators that have been found to facilitate or thwart desired behaviors.

Doctor-Patient Interaction

The doctor-patient dyad is an important relationship that impacts health outcomes directly and indirectly (see Figure 3). Communication between doctor and patient can lead to health outcomes either via proximal and intermediate outcomes, or directly.⁵⁸

⁵⁶ Lorber, M. F., Slep, A. M. S., Heyman, R. E., Xu, S., Dasanayake, A. P., & Wolff, M. S. (2014). Noxious family environments are associated with adult and childhood caries. *Journal of the American Dental Association*, *145*, 924-930.

⁵⁷ Halford, W. K., Petch, J., & Creedy, D. K. (2010). Promoting a positive transition to parenthood: A randomized clinical trial of couple relationship education. *Prevention Science*, *1*, 89-100.

⁵⁸ Street, R. L., Makoul, G., Arora, N. K., & Epstein, R. M. (2009). How does communication heal? Pathways linking clinician-patient communication to health outcomes. *Patient Education and Counseling*, *3*, 295-301.

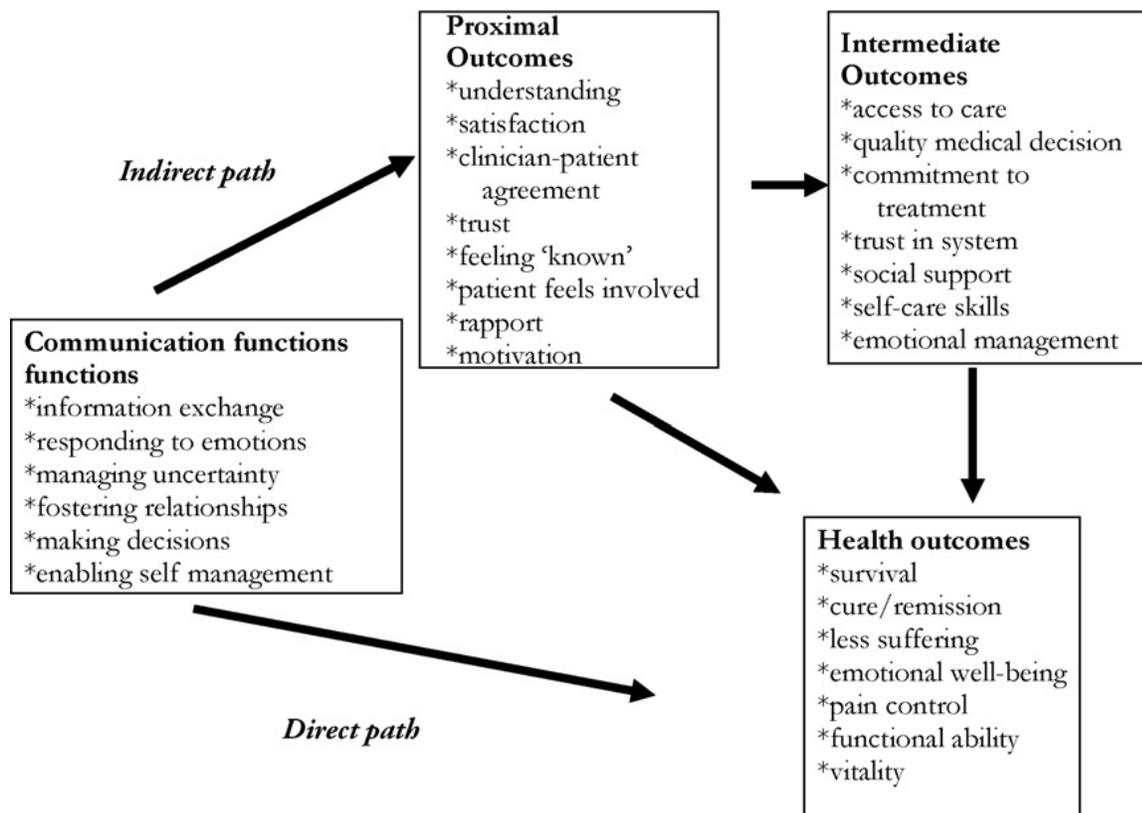


Figure 3: How doctor-patient communication impacts health.

The Roter Interaction Analysis System is commonly used in the field for assessing doctor-patient communication. It is similar in complexity and difficulty to first-generation couples observation systems. The capacity to make use of its 40 codes in a reliable way that could be translated into different providers' behavior is overwhelming. The essence of the coded behaviors has to be something smaller. A European consortium developed a simpler 4-code alternative that focuses on doctors' responses to patients' cues and concerns of unpleasant emotion.⁵⁹ Heyman and colleagues are currently piloting a study of doctor-patient interaction that is informed by couples observation techniques. They are examining the interplay between doctor and patient behavior and experienced emotion by using video-mediated recall of unpleasant emotion during a clinical visit.

A model of doctor-patient interaction involves the ecology of the individual patient (e.g., family, neighborhood, community, subculture, level of participation, health beliefs, adherence, distrust) and the health care system ecology of the clinician (e.g., practice/institution, health care specialty, health enterprise, subculture, cultural competency, perceived nonadherence, racial attitudes, engagement). Understanding the two ecologies and how they interact will

⁵⁹ Del Piccolo, L., De Haes, H., Heaven, C., Jansen, J., Verheul, W., Bensing, J., et al. (2011). Development of the Verona coding definitions of emotional sequences to code health providers' responses (VR-CoDES-P) to patient cues and concerns. *Patient Education and Counseling*, 82, 149-155. doi:10.1016/j.pec.2010.02.024.

further the understanding of the mechanisms by which doctor-patient interactions impact health outcomes.

Considerations for the Future

Behavioral observation is a flexible data collection method and yields high-quality information, but it is not without its challenges. Web-based applications and smartphones are making video collection easier (e.g., you can now record video via Skype). However, there are few opportunities for economies of scale in behavioral observation techniques: coding each observation from a large data set incurs the same level of burden as coding interactions from a pilot study. Proving the incremental validity of these methods is paramount, and yet it is still not very effectively done.

Investigators need to consider participant burden when planning and conducting intervention studies with tests of mechanisms, which inherently involve multiple measures and frequent rounds of data collection. Participants in Heyman's *Couple Care for Parents* intervention trial are pushing back in response to this burden. Recent literature has pointed to participant burden extending beyond the poverty of money and time, but to the "bandwidth cost of interventions," whereby well-meaning interventions can have disproportionate effects on individuals facing scarcity.⁶⁰ Heyman expressed concern that unless the frequent measures in intervention studies that also test mechanisms are quick and easy, the richer data that are being collected could come at the cost of participation. This is especially a concern for his study population of new parents who are facing several types of scarcity (e.g., time, sleep, money). New parents want help, but at a certain point, study participation can become overwhelming.

Interpersonal and Social Processes Group Discussion

Facilitators: Lisbeth Nielsen, PhD, National Institute on Aging; Melissa Riddle, PhD, National Institute of Dental and Craniofacial Research; Alison Miller, PhD, University of Michigan, Ann Arbor; Megan Moreno, MD, MPH, University of Washington

Throughout the presentations, several interpersonal and social processes were discussed as promising targets of behavioral change, as well as a sample of the available measurement approaches. Promising targets of behavior change along multiple levels of analysis—such as therapeutic alliance, parent-child alliance, couples alliance, hostility, rejection, attachment, loneliness, social connectedness, shared joy, and safety from threat—have to be malleable, measurable, and shown to alter a behavior of interest. Interpersonal and social processes address how behavioral decisions differ when a person is in a room alone versus with close friends versus in larger public contexts, and valid and reliable measures of the relevant constructs at each of those levels are needed. Participants held varying perspectives on the best strategies for coalescing the field of research to advance work on interpersonal and social processes targets for behavior change.

⁶⁰ Shan, A. K., Mullainathan, S., & Shafir, E. (2012). Some consequences of having too little. *Science*, 338, 682-685.

Developmental Perspective

Miller noted that assessments of interpersonal and social processes in children need to reliably measure the same constructs across developmental trajectories in a variety of contexts. The characteristics of the parent-child dyad and family bring unique contributions to each interaction. Assessments of these processes need to be conducted in a situation that allows the investigator to tap into the mechanism that is hypothesized to be the core construct, and then consider the individual difference factors that contribute to the situation. Lumeng added that a greater research focus is generally needed on interpersonal processes in the parent-child dyad to target behavior change in children.

New Technologies versus Human Interaction

Shinn's project demonstrated the impact of human accountability and therapeutic alliance in changing behavior and, as a result, impacting health outcomes. She noted that the interaction seemed to be about evoking a feeling of meaning in the patient that they were trying to change behavior (i.e., in this case, do painful swallowing exercises despite competing priorities) for others, such as the patient's family or the counselor. Heyman added that people are more likely to change behavior when it is meaningful and/or they perceive they are doing it for someone else. Loneliness has the opposite effect; a person has less incentive to change when he or she is lonely and not socially connected. Haushofer related this notion to recent data from Molly Crockett on how loss aversion is exacerbated when it is placed in the context of someone else. Individuals' willingness-to-pay was higher to avert harm to someone else than to themselves.

Participants wondered, however, whether and how new technologies could achieve similar behavioral results without as much human interaction. For example, a text message reminding a patient to take his or her medication or do daily exercises could be less intrusive and less expensive than human interaction. There are a number of variables involved that could be tested and measured to determine the combination of timing, packaging, ratio of technology contact versus human contact, phrasing, etc. that yields the most effective results in terms of behavior change.

There is excitement in the field of psychotherapy about the promise of text messaging. Many think that the rising generations will not require as much personal interaction because of their ease with technology and familiarity with connecting to others via social media. The question of whether the need for human interaction to evoke the same feelings of belonging and therefore changes in behavior is decreased for younger generations is an interesting one that has not yet been investigated experimentally.

Incremental Validity

Levenson noted that there is evidence and theoretical bases for the powerful role in behavior change of a number of social variables, such as therapeutic alliance or sense of being a team, couples dyads, loneliness, social connectedness, etc. Coding systems for behavioral observation of many of these constructs are time and resource intensive; however, they yield rich data. There needs to be a systematic exploration of the incremental validity for different types of measures of various constructs to ascertain when, for example, a measure such as self-report

or scales is sufficient for tapping the target of interest and when behavioral observation and coding adds a necessary richness.

Miller agreed that this is the question of measurement, but also noted that another important piece that needs to be considered is identifying at what level and where to locate an intervention—for example, changing fMRI response, social networks, parenting interactions, social norms—and that depends on the mechanism of interest.

Perspectives on Next Steps

As in earlier discussions, some participants stressed the importance of focusing on defining constructs that have the most evidence for being able to be targeted, measured, and intervened upon. Kumar suggested that what is most needed is an interdisciplinary group such as this to work on the development and validation of newly emerging measures rather than debate the merits of or cross-calibrate existing measures. The field will progress more rapidly if engineers develop measures jointly with people with interdisciplinary expertise to take advantage of new technologies to define constructs that are measurable.

Levenson discussed the challenge of moving this field forward. Currently, for any given intervention shown to be effective, researchers in different disciplines might investigate a hypothesized mechanism differently, targeting different constructs or targeting the same construct but with different measures. The challenge for moving forward is to take these disparate approaches and proceed in a more coordinated fashion. For example, if one person uses a scanner to measure neural activity and another codes facial expressions for the same underlying construct, then the different approaches need to contribute to a deeper and coordinated sense of what is going on. Right now, the science is not very incremental or progressive, to the detriment of building a cohesive field.

Day 1 Summary Group Discussion

Facilitators: Jonathan W. King, PhD, National Institute on Aging, and Melissa Riddle, PhD, National Institute of Dental and Craniofacial Research

The final discussion session of the first day included a spirited conversation with multiple viewpoints expressed. Some invited participants challenged the group to determine what it is about right now that puts the field in a position to move forward and what concretely can be accomplished in 5 years within the “science of behavior change,” which involves investigators from multiple disciplines using different approaches to study similar or overlapping constructs and how they affect different behaviors.

Strauman posited that the behavior change field appears to be at a stage where work is needed to match a population within a particular context to the most effective interventions. There is no lack of treatments for depression, for example, but what is needed is a finding or theory that explains what treatments work for whom in what contexts and through what mechanism.

Haushofer asserted that we know very little about the relative benefit of different types of interventions in terms of how much they change a given outcome and that interventions need to be compared against each other to make those determinations. How much can we change discounting behavior by intervening on stress? Is it better to intervene on the behavior itself, or intervene on targets such as stress or interpersonal relationship antecedents? What is needed is to identify effective interventions and then determine the mechanisms at work.

Scott Halpern posited that the NIH will have to indicate which approach is preferable: (1) Determine what interventions work for changing behavior and, along the way, learn something about the mechanisms of behavior change or (2) Study and test mechanisms of behavior change without necessarily clarifying whether particular mechanisms that target behavior have an impact on clinical endpoints. Riddle countered that both approaches are important and are not mutually exclusive. What is currently occurring is inefficient: interventions are tested one-at-a-time in separate ICs and with very little (and not at all systematic) attention to the mechanisms of behavior change. Supporting a more systematic, coordinated approach to behavior change research—both by encouraging the testing of mechanisms in ongoing intervention research and by systematically identifying, validating, and cross-calibrating targets that are intermediate processes for multiple behaviors and clinical endpoints—is the goal of a future SOBC Program.

Gotlib suggested a systematic approach at a smaller scale for the 5-year program to have an impact on the field: identify a behavior to change, identify three mechanisms hypothesized to be at play for the selected behavior, and then determine how the mechanisms will be targeted, what the possible moderators are, and the conditions under which they will work. King noted that this is precisely the kind of systematic work that the group envisioned would be supported in the future.

Levenson summarized several pathways forward: (1) identify hypothesized mechanisms of behavior change and test them; (2) cross-calibrate among different measures of the same construct; and (3) test putative mechanisms of observed changes in behavior. Once the mechanisms of behavior change are better known, more complicated interventions can be developed that use the proven mechanisms. King noted that the Common Fund has supported work in the past that is more systematic and programmatic if it is what is needed to move the field forward and it is not something that an individual IC would be likely to support.

Riddle, King, and Nielsen reiterated the vision that at the end of another 5-year period, the program will have supported work to systematically study and test a sample of promising intermediate-level mechanisms along the pathway to multiple behaviors that impact a broader range of clinical endpoints and, in doing so, provide a proof-of-concept model for incorporating the testing of mechanisms into behavior change research and produce validated and cross-calibrated measures for the constructs involved. A long-term goal is to change the behavior of the NIH and investigators by setting up the expectation that behavior change research will include the testing of mechanisms by providing measures, methods, and a model for doing so.

Creative Destruction of the Science of Behavior Change

William Riley, PhD, Acting Director, Office of Behavioral and Social Sciences Research, and Chief, Science of Research and Technology Branch, National Cancer Institute

Riley acknowledged the strong support and leadership that the SOBC Program receives from its co-chairs: Patricia Grady, Director of the National Institute of Nursing Research (NINR), Richard Hodes, Director of the NIA, and Richard Suzman, Director of the Division of Behavioral and Social Research at the NIA.

A typical meta-analysis of a behavior change intervention results in a modest to moderate effect and yields a fair amount of heterogeneity between studies involving multi-component interventions. The results of a typical trial do not yield a cumulative sense of what pieces of the intervention worked and why. Topol's book on the creative destruction of medicine discusses the impact of the digital revolution and rapid technological advancement to biomedical research;⁶¹ however, these changes also apply to behavioral science. In late 17th-century England, it was observed "[N]early all the grandest discoveries of science have been but the rewards of accurate measurement."⁶²

Riley reviewed three types of technological advances in behavioral measurement: Item Response Theory (IRT) and Computer Adaptive Testing, ecological momentary assessment (EMA), and passive sensor technologies.

Item Response Theory

The most common and recent example of advances in IRT is the Patient Reported Outcomes Measurement Information System (PROMIS), which is another Common Fund program.⁶³ IRT is not new or novel; the education field has been using it since the 1960s, but its application to behavioral science is more recent. The basis of IRT is that for every response to a given question there is a probability curve associated with where an individual falls in the underlying latent trait that the question is trying to tap. Typically in classical measurement development, all items are treated the same even though they should have different weights. IRT allows for this type of weighting. The implication of IRT for Computer Assisted Testing is that later questions can narrow in on the construct more quickly based on responses to previous questions.

IRT also enables users to calibrate existing measures on the same scale. What is needed in behavioral science is not consensus measures, but consensus constructs and a consensus *metric* for measures of those constructs. For example, blood pressure can be measured with multiple tools (e.g., sphygmomanometer and stethoscope, electronic cuff), yet what is constant is the

⁶¹ Topol, E. (2012). *The Creative Destruction of Medicine: How the Digital Revolution Will Create Better Health Care*. New York: Basic Books.

⁶² *The Athenaeum: Journal of English and Foreign Literature, Science, the Fine Arts, Music and the Drama*. (1871). No. 2284, p. 179.

⁶³ More information about PROMIS can be found at <http://www.nihpromis.org>.

metric of measurement and the way it is reported: systolic over diastolic millimeters of mercury. There are multiple measures of depression, and every instrument has a different scale, rendering them incomparable. Seung Choi (Northwestern University) and others have done work to calibrate different depression measures on the same scale.

Ecological Momentary Assessment

EMA allows for experience sampling using technology such as text messaging. This method is better than self-report measures for capturing accurate data in the moment. EMA can involve single or repeated measurements on flexible fixed, random, or mixed schedules. A smartphone is a ubiquitous device that can allow for participant sign-up and management, intake and follow-up survey links, text message experience sampling, and response data monitoring.⁶⁴

Passive Sensor Technologies

Physical activity is implicated in almost every chronic disease, but measuring physical activity at the population level is a challenge. New technological advances in wearable sensors and the availability of wireless accelerometers are enabling population-scale measurement of physical activity. There are now miniature, low-cost devices that measure human motion using re-designed accelerometers in a user-friendly format. The Wocket, for example, allows for a fine-grained assessment of activity that is economically feasible for population-level studies. It streams data in real time to a mobile phone. Sensors that are worn under clothing 24/7 are miniature, thin, and ergonomic, and the smartphone is held or carried as usual. Pattern recognition algorithms running continuously on the phone detect physical activities in real time. Emerging technologies and assays for adherence monitoring include the Xhale SMART breathalyzer for GRAS drug taggants, drug concentrations via hair samples or dried blood spots, Proteus pill microchips and sensor (tracks whether the person took the pill), and GlowCaps (tracks whether the person opened the cap of the medicine). Another example is Kumar's work on the Autosense system, which includes armband (e.g., alcohol, temperature, GSR, accelerometer) and chestband sensors (e.g., ECG, respiration, GSR, ambient and skin temperature, accelerometer) that communicate wirelessly with an Android smartphone.

Through work supported by the GEI Exposure Biology Program at the National Institute of Environmental Health Sciences (NIEHS), NJ Tao has developed a sensor system that measures volatile aromatic hydrocarbons and acidic vapors with the level of sensitivity needed for environmental epidemiological studies. This technology will enable the assessment of individual exposure levels during a daily routine, which is superior to the current method of averaging exposures over a geographical region.

As exciting as developments in passive sensor technologies are, it is still important to define common constructs and metrics.

⁶⁴ Ginexi, E. M., Riley, W., Atienza, A. A., & Mabry, P. L. The promise of intensive longitudinal data capture for behavioral health research. *Nicotine and Tobacco Research*, 16, S73-S75.

Behavioral Data

Technological advancement is supporting new modalities for prospective data collection as well, including citizen science and crowdsourcing approaches and opt-in internet panels. Behavioral data gleaned from consumer-based data sources such as social media sites, Google internet searches, cell phone use, cable box data, and car black box data could be used to mine data about behaviors.

The Big Data to Knowledge (BD2K) program, also funded by the Common Fund, seeks to create new approaches, methods, tools, and software to enable biomedical scientists to capitalize more fully on the large, complex, and diverse data sets being generated by researchers. The field is moving from a focus on just “long” data sets with large sample sizes to “wide” data sets that include longitudinal data that track behavior and behavior change within a person over time. For example, applying pattern recognition analytics to all available biopsy data on breast cancer biopsies actually predicts features associated with survival outcomes better than pathologists.⁶⁵

Grantee Updates: Session 4

Acceptance, Efficacy, and Effectiveness of Incentives for Smoking Cessation

Scott D. Halpern, MD, PhD, University of Pennsylvania

This project is a collaboration with David Asch and Kevin Volpp and others with support from SOBC, NIA, and in-kind contributions from CVS/Caremark. The project involves a five-arm smoking cessation RCT with CVS employees (see Table 3). Investigators randomized 2,538 patients to one of the arms, and 1,528 accepted their assigned intervention arm. This is the largest randomized trial of incentives to date. The primary outcome is sustained abstinence for 6 months. Secondary outcomes include abstinence for 14 days, 30 days, and 12 months.

The five-arm RCT includes usual care and four treatment groups. The usual care group received access to behavioral counseling and nicotine replacement. The intervention arms included combinations of financial framing (gain or loss) and individual versus group rewards for performance. The expected value of all programs was \$800 over the course of the study. Enrollment is done on a rolling basis, and groups are formed based on quit date. The collaborative reward group receives more money as more people are successful, and these group members are linked via chat rooms. Members of competitive groups receive less reward as more people succeed and do not interact with each other.

⁶⁵ Beck, A. H. et al. (2011). Systematic analysis of breast cancer morphology uncovers stromal features associated with survival. *Science Translational Medicine*, 3, 108-113.

| Table 3. Incentive Structures to be Compared (arms 2 – 5) | | |
|---|--|---------------------------|
| Financial Frame | Payments based on performance of: | |
| | Individual | Group |
| Gain | (2) Individual Rewards | (4) Collaborative Rewards |
| Loss | (3) Individual Deposits | (5) Competitive Deposits |

In the past year, the team met recruitment goals, completed data collection, and completed analyses of three main effects: (1) acceptance of different incentive programs among those assigned; (2) efficacy (i.e., met primary endpoint of not smoking for 6 months) conditional on accepting; and (3) effectiveness (i.e., how many of those assigned achieved the primary aim as a product of acceptance and success).

To be effective, incentive programs need to be acceptable to targeted participants.⁶⁶ Therefore, effectiveness equals acceptance multiplied by efficacy. Patients had significantly higher acceptance rates of the individual reward (95 percent probability of acceptance) and collaborative reward groups (85 percent probability of acceptance) assignments versus the individual deposit (13 percent probability of acceptance) and competitive deposit groups (15 percent probability of acceptance). The deposit groups required a \$150 deposit from each participant at the beginning of the program.

The 6-month abstinence rates of all participants were 6 percent in usual care, 15 percent in rewards groups, and about 10 percent in the deposit groups. Halpern noted that for smoking cessation, there is no other treatment that comes close to a 15 percent success rate. Because acceptance of the deposit groups was so low, the efficacy of the deposit groups conditional on acceptance is actually much higher at 60 percent. However, the people who accepted the deposit contract groups are different than others in the sample, and their characteristics will be examined in post-analysis comparisons. Abstinence rates at 6 months were not statistically different between group-based incentive-based arms and the individual-based incentive arms. However, the group-based incentives result in higher abstinence rates at 12 months: 7.5 percent in group-based incentives versus 5.4 percent in the individual-based groups ($p = 0.05$).

Given that randomization does not imply acceptance of the assigned intervention or compliance with the treatment, a better comparison is to use instrumental variable analysis. Instrumental variable and principal stratification methods were used to overcome the biases typically associated with per-protocol analyses. The randomization arm (i.e., the initial treatment assignment) is treated as the instrument in the instrumental variable analysis. As a result, the estimated treatment effect of each intervention on smoking cessation is adjusted by the percentage of assigned patients who accepted the treatment arm. This differs from traditional intention-to-treat analysis that only looks at the treatment effect based only on

⁶⁶ Halpern, S. D., Asch, D. A., & Volpp, K. G. (2012). Commitment contracts as a way to health. *BMJ*, e522.

assignment and not acceptance. In addition, the instrumental variable approach uses the data on all randomized patients, rather than only on those who accept the intervention, and then adjusts for acceptance, which attenuates the selection effects.

The instrumental variable analysis found that the efficacy of rewards groups among participants who would accept a reward intervention is 10.7 percent greater than usual care. The efficacy of deposit groups among participants who would accept a deposit intervention is 30.8 percent greater than for usual care. The best estimate of the efficacy of deposit versus rewards groups among participants who would accept either intervention (assuming that the underlying odds of quitting among participants who would accept deposits is greater than the odds of quitting among participants who would only accept rewards) is 13.2 percent.

Smoking employees cost employers roughly \$2,739—5,816 more per year than nonsmoking employees. The cost-per-quit for this program was on average about \$500.

In summary:

- The RCT shows that financial incentives triple quit rates at 6 months versus free counseling and nicotine replacement therapy among all employees to whom the programs are offered.
- Incentive programs requiring \$150 deposits are unattractive to many employees, particularly low-income employees. However, for those who accept deposit programs, the programs are at least twice as efficacious as pure reward-based programs.
- Together, these findings suggest that programs combining smaller or even “virtual” deposits with generous matches will be optimally effective across populations of employees. The key will be to find the right balance between acceptance and efficacy.

A future trial will use an opt-out design rather than screen for individuals who indicate they have thought about smoking cessation. This approach will likely reduce the success rates but will allow investigators to examine how the intervention would work in the real world.

Self-Regulation, Health, and the Health-Education Gradient

Henry Saffer, PhD, National Bureau of Economic Research

Saffer’s project is an extension of earlier work on the different effects of alcohol advertising and price on alcohol use, which pointed to the need to further investigate the role of self-regulation. The goals of this project are to measure self-regulation and its effect on health choices, investigate whether an individual’s self-regulation differs by health choice, and estimate the effect of self-regulation on the health-education gradient. For this study, self-regulation is defined as an individual’s ability to defer an immediate reward for a future reward, and it can differ by outcome. The education gradient refers to the effect of education on health choices. Saffer hypothesized that self-regulation has a moderating effect on the effectiveness of policy variables (i.e., people with higher or lower ability to defer immediate rewards are more or less affected by advertising and price).

This project uses data from the National Longitudinal Study of Adolescent Health (Add Health), which is comprised of a nationally representative sample of individuals in grades 7-12 in the United States during the 1994-1995 school year. The cohort has been followed into young adulthood with four in-home interviews, the most recent being in 2008. Data include respondents' social, economic, psychological, and physical well-being and contextual data on the family, neighborhood, community, school, friendships, peer groups, and romantic relationships. The working data set used in this analysis was created with data from waves 3 and 4 of Add Health and included a sample of about 13,000 individuals.

Self-regulation, as defined for the purpose of this secondary data analysis, is a latent variable, and proxy variables can be used to approximate the latent variable. A novel empirical approach using the Lubotsky-Wittenberg method minimizes the total error associated with the proxy variables. This method provides a set of weights for each proxy that maximizes total information relative to methods, which only extract a common variance from the proxies. The Lubotsky-Wittenberg method does not restrict the self-regulation variable to be the same for each outcome.

Saffer constructed a set of regressions using proxy variables and other independent variables on several dependent outcomes. Proxy variables were constructed for self-regulation (based on 20 Add Health items); genetics, maternal attachment, conscientiousness scale, stress scale, anxiety scale, religious adherence, and attention deficit hyperactivity disorder (ADHD). Other independent variables included education, gender, age, race/Hispanic origin, income, marital status, number of children, and time and state fixed effects. Outcomes included being a smoker, cigarette consumption of smokers, alcohol consumption of drinkers, binge drinker, marijuana use in past 30 days, number of drug abuse symptoms, ever been arrested, index of criminal behavior, gambler, obese, and BMI.

Analyses indicate that the self-regulation proxy and the ADHD proxy were statistically significantly associated with being a smoker, cigarette consumption, drinking alcohol, being a binge drinker, marijuana use, drug abuse, ever arrested, crime index, being a gambler, being obese, and BMI. The religious adherence proxy was associated with all the same outcomes with the exception of crime index and being a gambler. The estimated self-regulation proxy was highly correlated with the other outcomes, with the exception of self-regulation for obesity and BMI.

Individual heterogeneity in self-regulation may limit the effectiveness of traditional public policy for health-related outcomes and on obesity in particular. Further research on the interaction of price, advertising, and other policies with self-regulation is needed to examine the potential limits of policies related to health choices.

Lunch is in the Bag

Courtney Byrd-Williams, PhD, University of Texas, Austin

Byrd-Williams and Principal Investigator Deanna Hoelscher received an administrative supplement from the SOBC Program for an R01 grant from NCI to test the Lunch is in the Bag intervention intended to increase fruit, vegetable, and whole grains packed in lunches and consumed by preschoolers. The R01 project involves a 6-week, multi-component intervention consisting of parent handouts and classroom-based activities and books. The RCT randomized 30 childcare centers to the intervention or a waitlist in two waves and two cohorts. About 40 percent of the childcare centers ask parents to provide lunch, and packed lunches tend to be less healthy than lunches provided by the center: packed lunches do not meet guidelines for fruits and vegetables, 93 percent do not contain the recommended amounts of fiber, and packed lunches include alarmingly high levels of sodium (114 percent of daily recommended intake).

Investigators conducted debriefing discussions with teachers and staff at the childcare centers and determined that teacher support and behaviors may be an important target for increasing what the parents pack in lunches. To test this, investigators developed enhanced teacher training to help parents pack and preschoolers to eat more fruits, vegetables, and whole grains. The enhanced teacher training included nutritional knowledge, best practices for structuring mealtime, and how to provide support and reinforcement for parents.

The purpose of the SOBC Program supplement is to examine the effect of enhanced training for teachers on supporting healthful eating and improving the lunches parents pack for preschoolers. Twelve childcare centers were randomized to one of four conditions: (1) Lunch is in the Bag program only; (2) Lunch is in the Bag plus enhanced teacher training; (3) enhanced teacher training only; and (4) waitlist control group. The primary outcome is the number of servings of fruits, vegetables, and whole grains packed in lunches *and* consumed by children. The supplement sought to develop and test the effects of teacher training beyond familiarizing the teachers with the materials. Feedback from parents and teachers indicate that teacher reinforcement of healthful lunches increases child requests for parents to pack a healthful lunch. Teachers requested nutrition information and suggestions for how to talk to children and parents about nutrition.

To date, the intervention has been implemented in the treatment centers. One-hour teacher training has been delivered at six childcare centers with 91 percent fidelity. Training evaluations showed significant increases in teacher nutrition knowledge and self-efficacy to influence children's eating behaviors. The team has entered and cleaned the observation and survey data, and analysis is ongoing.

Based on teacher self-report data, preliminary analyses indicate that teachers in the two intervention groups that included enhanced teacher training reported higher levels of talking to parents about what their child ate that day, writing notes to parents that praised them for packing a healthful lunch, and praising children whose parents packed fruits, vegetables, and

whole grains. Amounts of whole grains and vegetables packed increased, but not fruits, mostly because of ceiling effects (baseline lunches tended to already have fruit).

The 1-hour teacher training affected teachers' supportive behaviors as measured by self-report, which suggests that the enhanced teacher training is a promising addition to the Lunch is in the Bag program to improve packed lunches. Although there is evidence of changes in parents packing lunches in the enhanced teacher training group, child consumption is not affected. Next steps will include creating a series of trainings to increase teacher support and influence children's consumption. Center directors reported that teachers do not have time for more training, but the teachers themselves requested more.

The team also collected information about the childcare center context, including center snack policies and whether the center or parent provides snacks. The center directors might be appropriate targets for intervention as well to change the culture of the nutritional environment.

Environmental Measures as Moderators

Facilitators: Jonathan W. King, PhD, National Institute on Aging, and Scott D. Halpern, MD, PhD, University of Pennsylvania

Behavioral Economics and Environmental Manipulations

Halpern provided introductory comments from the behavioral economics perspective on the choice environment. Much of behavior is about choice. Behavioral economics seeks to identify ways to manipulate the choice environment, which then impacts health outcomes. What currently holds promise is the idea that the choice environment can be manipulated to be preferentially advantageous for people who differ in characteristics at baseline, such as temporal discounting.

Environmental interventions can be placed along a hierarchy of invasiveness from heavy- to light-handedness: for example, prohibiting certain choices completely (e.g., not hiring people who self-report as smokers), defaulting people's choices (e.g., opt-out retirement contributions), using financial incentives or disincentives, or staging the environment to encourage and/or automate desired choices (e.g., cookies in the back, fruit in the front of cafeteria line; fly in the urinal). Heavier-handed approaches are presumed to be more effective, but all things being equal we would want to choose things lower on the hierarchy because they pose a lesser threat to autonomous choice. In addition to experimental manipulations designed to assess light- and heavy-handed approaches to environmental manipulation and its effects on individuals' behavior, there might be opportunities for natural experiments such as characterizing existing neighborhoods and then examining how people respond to their environments differently.

Person-by-Cue-by-Strategy-by-Context

In the final year of Halpern's project, the investigators are interested in examining interactions between differences in baseline performance and characteristics of individuals and the assigned intervention arms. For example, people with steep delay discounting functions are those who have the most difficult time quitting smoking because the health benefits in the future are discounted disproportionately compared to others; therefore, one type of incentive arm may be more effective for people with this characteristic. This idea is related to what Ochsner, Strauman, and others noted earlier: the 2.0 of behavioral science is a personalized model that accounts for unique person-by-cue-by-strategy-by-context interactions. The statistical challenge with this approach is that the sample sizes required to detect such interactions are 50 percent higher than the size needed to detect a main effect. With the advances in digital technology and remote monitoring, larger trials might be feasible, and this approach could be used to identify mechanisms.

Ochsner shared his perspectives on measuring context. In emotion regulation research, interpersonal context includes dispositional variables such as gender, age, developmental history, and extrapersonal context, which refers to contextual determiners of when certain strategies are effective or not—does the contextual feature increase or decrease the potency of the strategy. Extrapersonal contextual variables could include available choices, stressors in the environment, or moderators. Self-control is not a singular thing. Individuals will exert self-control in some contexts and not in others. The cognitive and affective neuroscience fields naively thought people with weak prefrontal cortices would have generalized deficits and be bad at all decisions, but this has not been found to be the case. Gotlib agreed and related these ideas to the field of psychotherapy: the question is not “does psychotherapy work,” but “for whom, and under what conditions does psychotherapy work.”

Developing Partnerships

Halpern noted that his research group has expertise in developing partnerships with real-world settings such as large corporations, health systems, and insurers where interventions can be tested at scale. However, they do not necessarily have expertise in defining and refining constructs or developing consensus metrics or measures. Halpern's work is focused on determining what intervention works best, but these large-scale studies present a tremendous opportunity to learn about the underlying mechanisms at work, often in real-world populations. There could be a better structure in place to facilitate a marriage of intervention researchers and researchers working on specific mechanisms. Nielsen noted that there is not currently a widespread way to promote the skill set necessary to develop successful partnerships such as Halpern and a handful of other researchers have been able to do. Weber offered an example strategy that might be useful for SOBC 2: the Coordinating Center for the Common Fund's Health Care Systems (HCS) Research Collaboratory is tasked with compiling and disseminating best practices and lessons learned for successful stakeholder engagement.

Closing Group Discussion

King summarized the planned goals for SOBC 2. The new program would support development of a subset of promising consensus constructs and consensus metrics for targets in a small number of teams (perhaps three to six) that would include those interested in the constructs themselves and also psychometricians, interventionists, clinicians, and others. The constructs and measures need to be appropriate for the clinical environment and relevant to specific behavior change outcomes. SOBC 2 will not support large clinical trials, but rather the target identification, assay development, target validation, and the introduction of selected assays into ongoing trials.

Nielsen added that we need to determine how best to move the field forward in 5 years by better defining constructs, providing tools and measures and guidance on how to test mechanisms in intervention studies, and incentivizing the type of work needed. Fractioning the entire domain of cognitive control is beyond what is being suggested here. Even getting investigators oriented to a set of constructs they should be measuring in ongoing work would be of high value.

Readiness to Act

Riddle reiterated the goal of compiling consensus constructs and consensus metrics across levels of analysis and disease entities, at least for a subset of exemplar targets, and inquired whether or not the field is ready for that step. There needs to be some general agreement on a reference measure to develop or define a consensus metric or reference scale. For example, the work done on calibrating depression measures was based on 20 different self-report measures of well known and agreed upon constructs that had been examined in thousands of studies.

Nielsen noted that there are many self-report measures of self-control that could be evaluated to extract commonalities and the extent to which they are measuring the same construct. However, it is not yet a coherent effort. Also, there are others assessing the same or similar constructs of self-control at different levels of analysis—at the neurobiological, psychological, and behavioral levels. There is an array of tasks and an array of operational definitions of the construct. What is missing is a cross-disciplinary group of researchers convening to test and define to what extent those measures or tasks capture the same underlying construct.

Strauman felt optimistic that this field is rich in theory that can guide ideas about which interventions work best for which individuals. The question of why they work and by what mechanism still needs to be studied, but the field is awash in good theory and good constructs. What are needed are consensus constructs and good a priori hypotheses.

Levenson stressed that a concept such as “self-control” is an order of magnitude more complex than the concept of “depression.” A person’s ability to regulate hunger is likely not the same as

the ability to regulate anxiety or seek a long-term goal. Before a translational mechanism can be identified, similar types of control would need to be grouped together.

Ochsner reiterated that terms such as self-regulation or self-control are so broad at a descriptive behavioral level and represent a class of constructs that might be pursued, but they are not constructs themselves. Any number of psychological constructs might work together to be described at a behavioral level as exhibiting self-control. It is an umbrella term. Classes of behavioral strategies are similar in terms of what you are asking people to do and might depend on what appear to be the same underlying circuitry or region critical for carrying out tasks; however, the current technology of brain mapping is not advanced enough to enable identification of careful temporal patterning. Today's brain mapping techniques allow for crude comparisons akin to comparing two qualitatively different types of music with only the knowledge of what instruments are playing. fMRI is not sensitive enough to determine temporal patterning. Better instrumentation is needed and will come with technological advancement. Phelps added that one type of research that has not been done to examine areas of the brain implicated in multiple constructs is to systematically examine individual high-resolution images of the brain to identify, for example, the breakdown of different types of self-control and how they differ in the lateral prefrontal cortex. There is a little of this work in monkeys on value versus inhibition, but because this type of work in humans is not innovative or hypothesis driven it tends not to be funded.

Clarity of Language

Phelps noted that the distinction between cognitive control and emotion regulation is an example of a distinction of ideas that are not completely unrelated but have different goals in terms of how they change behavior. She expressed concern that researchers in different fields do not use the same terms and might not understand what the umbrella term self-regulation includes. It will be more interesting to focus on how different components of self-regulation can be utilized to understand behavior and behavior change and enhanced to intervene on behavior change. Strauman noted that a publication by Todd Heatherton (Dartmouth) about cognitive neuroscience and self-regulation distinguished among different aspects of self-regulation including goal pursuit and affect regulation and discussed that the public health relevance might be a useful consensus point at which to start.

Defining What Works

Halpern cautioned that defining "what works" should be very clearly qualified by what works within whom, to what degree, and under what conditions. The pieces that "work" can be used as the comparison intervention so that more innovative interventions can be developed based on theories and basic science evidence of mechanisms. Improvements to existing interventions should be predicated on theory.

Haushofer commented that there is a distinction between starting with a desired outcome, finding effective interventions for that outcome, and along the way using that work as a point of departure for learning about mechanisms versus starting with consensus constructs and mechanisms and using them to generate hypotheses about what interventions will work.

Others labeled this the top-down (i.e., starting with interventions and decomposing them to find the elements responsible for effects) versus bottom-up (i.e., identify causative mechanisms to include in intervention development) approach.

Haushofer and Shinn expressed a preference for focusing on interventions that work because they do a pretty good job of constraining the constructs and mechanisms we should be thinking about, and then examining the mechanisms that underlie the way the interventions work. Shinn believed that focusing on an intervention that improves a situation is a better use of limited time and research funding than chasing a particular target that may or may not pan out. Lumeng acknowledged that in her role, she sits between the “top-down, bottom-up” research worlds. As a clinician she wants to simply provide effective interventions to patients. Yet the lack of truly effective interventions for obesity frustrates her and compels her to do more bottom-up work to inform the development of more efficacious interventions.

Halpern admitted that the expressed dichotomy between top-down and bottom-up research was not resonating with him. There is a finite number of research dollars, and we all want to reduce stress, cardiovascular events, and obesity, for example. Funds should not be spent on interventions that target health-related endpoints that do not build upon what has been learned about the mechanisms by which stress can be reduced, for example. Research on constructs and mechanisms will inform better-designed interventions. These approaches are not mutually exclusive.

Several agreed that both approaches are necessary and doable. The SOBC 2 proposal includes goals for identifying constructs, measures, and targets as well as supporting the inclusion of testing mechanisms within ongoing trials. The SOBC Expert Review Panel recommendation of creating a registry of trial characteristics and results to attempt to identify what has already been tested in terms of the underlying mechanisms is also included in the new proposal.

Riddle reiterated the findings of the behavioral intervention portfolio analysis conducted by the SOBC Working Group: few projects incorporated tests of mechanisms, even when the mechanisms hypothesized to be responsible for behavior change were explicitly stated in the application. This does not mean that the intervention was not built on the assumptions of mechanisms or a thoughtful theoretical foundation. But what is *not* being routinely included is a *test* of the mechanism. It cannot be enough to say “I think the reason the intervention will work is X mechanism.” SOBC 2 is seeking to spur a paradigm shift in NIH-funded behavior change research such that the majority of future applications instead say “I think the reason the intervention will work is X mechanism, and I am going to test whether we moved X and then whether that moved the outcome.”

Gotlib noted that this effort appears to be analogous to the RDoC initiative. NIMH FOAs do not focus on clinical trials for a particular disorder; rather, they require that an intervention target (one of the RDoC constructs) is included as the mechanism of change. RDoC has done a lot of the footwork in constructing this system from the perspective of mental health. If SOBC could

attain some agreement on what are the short list of phenotypes, mechanisms, or constructs of interest, then funding opportunities could be framed in the same way.

One Target, Multiple Outcomes

Levenson offered two promising candidates for common behavioral factors with reasonable evidentiary basis that underlie a person's trajectory to healthful or unhealthful living are conscientiousness and the ability to self-regulate. These processes could be broken down into very small proximal units to retrain in simple things such as attention and signal detection as a way of shifting them onto more complex dimensions. An intervention to increase conscientiousness based on careful understanding of the construct and how to impact it, for example, down stream would help people prevent overeating, substance abuse, etc. These could be the behavioral analogs to inflammation, as Slavich noted, as a common factor explaining multiple problems.

Incentives to Collaborate

Throughout the discussions, it was made clear that this effort will require collaboration among researchers from multiple disciplines who work at a variety of levels of analysis across a continuum of settings from the laboratory to the clinic. This range of expertise and experience is represented on the SOBC Working Group as well. SOBC 2 funding opportunities do have the capacity to prompt changes in investigator behavior. Moreno used the analogy of trying to get people to use public transportation: individuals know public transportation is better for the environment, but they still want their own car. SOBC FOAs can incentivize investigators to simply carpool with others with whom they have not shared a ride before—prioritize funding for interdisciplinary teams, force investigators to attend these types of meetings and work together more closely, and bring postdoctoral students and collaborators to spread the reach—and SOBC 2 could have a very large impact on investigator behavior and the expectations for NIH-funded behavior change research.

APPENDIX A: MEETING AGENDA

| Monday, June 23, 2014 | | NIH, Building 31, Room 6C6 |
|-----------------------|--|--|
| 9:00 a.m. | Welcome and Introductions | Jonathan W. King |
| 9:10 a.m. | Current Perspectives on Behavior Change: Insights from the SOBC Expert Review Panel | Michael Otto |
| 9:30 a.m. | The Future of SOBC | Jonathan W. King |
| 9:45 a.m. | GRANTEE RESEARCH PRESENTATIONS 5 minutes each plus Q&A | Barbara Fredrickson Timothy Strauman Ian Gotlib Kevin Ochsner |
| 10:15 a.m. | BREAK | |
| 10:35 a.m. | Self-Regulation Measures and Targets <i>Facilitators:</i> Lisbeth Nielsen, William Klein, Timothy Strauman, and Kevin Ochsner <i>Introductory Remarks (15 minutes)</i> <i>Roundtable Discussion (60 minutes)</i> | |
| 11:50 a.m. | LUNCH | Building 31 Cafeteria |
| 12:50 p.m. | GRANTEE RESEARCH PRESENTATIONS 5 minutes each plus Q&A | Elizabeth Phelps Julie Lumeng and Alison Miller Johannes Haushofer |
| 1:10 p.m. | Stress and Stress Reactivity Measures and Targets <i>Facilitators:</i> Susan Czajkowski, Wendy Weber, and Elizabeth Phelps <i>Introductory Presentations (15 minutes each)</i> <i>Roundtable Discussion (60 minutes)</i> | Santosh Kumar George Slavich |
| 2:40 p.m. | BREAK | |
| 3:00 p.m. | GRANTEE RESEARCH PRESENTATIONS 5 minutes each plus Q&A | Megan Moreno Meike Bartels Eileen Shinn |

APPENDIX B: PARTICIPANT LIST

GRANTEES AND INVITED PRESENTERS

Meike Bartels

VU University Amsterdam
Email: m.bartels@vu.nl

Courtney Byrd-Williams

University of Texas, Austin
Email: courtney.e.byrdwilliams@uth.tmc.edu

Barbara Fredrickson

University of North Carolina at Chapel Hill
Email: blf@unc.edu

Ian Gotlib

Stanford University
Email: ian.gotlib@stanford.edu

Santosh Kumar

University of Memphis
Email: skumar4@memphis.edu

Scott Halpern

University of Pennsylvania
Email: shalpern@exchange.upenn.edu

Johannes Haushofer

Massachusetts Institute of Technology
Email: Haushofer@gmail.com

Richard Heyman

New York University
Email: richard.heyman@nyu.edu

Robert Levenson

University of California, Berkeley
Email: boblev@socrates.berkeley.edu

Julie Lumeng

University of Michigan, Ann Arbor
Email: jlumeng@umich.edu

Alison Miller

University of Michigan, Ann Arbor
Email: alimill@umich.edu

Megan Moreno

University of Washington
Email: megan.moreno@seattlechildrens.org

Kevin Ochsner

Columbia University
Email: ko2132@columbia.edu

Michael W. Otto

Boston University
Email: mwotto@bu.edu

Elizabeth Phelps

New York University
Email: liz.phelps@nyu.edu

Henry Saffer

National Bureau of Economic Research
Email: hsaffer@gc.cuny.edu

Eileen Shinn

University of Texas, MD Anderson Cancer
Center
Email: eshinn@mdanderson.org

George Slavich

University of California, Los Angeles
Email: gslavich@mednet.ucla.edu

Timothy Strauman

Duke University
Email: tjstraum@duke.edu

PARTICIPANTS

Dorothy Castille

National Institute on Minority Health and
Health Disparities
Email: dorothy.castille@nih.gov

Elaine Collier

National Center for Advancing Translational
Sciences
Email: ec5x@nih.gov

Susan Czajkowski

National Heart, Lung, and Blood Institute
Email: czajkows@mail.nih.gov

William Elwood

Office of the Director
National Institutes of Health
Email: elwoodwi@od.nih.gov

Holly Garriock

National Institute of Mental Health
Email: holly.garriock@nih.gov

Lynne Haverkos

Eunice Kennedy Shriver National Institute of
Child Health and Human Development
Email: lynne.haverkos@nih.gov

Lenwood Hayman

University of Michigan-Flint
Email: lwhayman@umich.edu

Jonathan W. King

National Institute on Aging
Email: kingjo@nia.nih.gov

William Klein

National Cancer Institute
Email: kleinwm@mail.nih.gov

Minda Lynch

National Institute on Drug Abuse
Email: minda.lynch@nih.gov

George Niederehe

National Institute of Mental Health
Email: gniedere@mail.nih.gov

Lisbeth Nielsen

National Institute on Aging
Email: nielsenli@nia.nih.gov

Wendy Nilsen

Office of the Director
National Institutes of Health
Email: nilsenwj@od.nih.gov

Lisa Onken

National Institute on Drug Abuse
Email: lisa_onken@nih.gov

Frank Perna

National Cancer Institute
Email: pernafm@mail.nih.gov

Mary Perry

Office of the Director
National Institutes of Health
Email: perryrna@mail.nih.gov

Melissa Riddle

National Institute of Dental and Craniofacial
Research
Email: riddleme@mail.nih.gov

William Riley

Office of the Director
National Institutes of Health
Email: wiriley@mail.nih.gov

Usha Sharma

National Institute of Allergy and Infectious
Diseases
Email: usha.sharma@nih.gov

David Shurtleff

National Center for Complementary and
Alternative Medicine
Email: dshurtle@mail.nih.gov

Janine Simmons

National Institute of Mental Health
Email: simmonsj@mail.nih.gov

Wendy Weber

National Center for Complementary and
Alternative Medicine
Email: weberwj@mail.nih.gov

Lois Tully

National Institute of Nursing Research
Email: lois.tully@nih.gov

Elizabeth Wilder

Office of the Director
National Institutes of Health
Email: elizabeth.wilder@nih.gov

NIH CONTRACTOR STAFF

Chandra Keller-Allen

Rose Li and Associates, Inc.
Email: chandra.keller-allen@nih.gov

Rose Maria Li

Rose Li and Associates, Inc.
Email: rose.li@roseliassociates.com

Samantha Lee

Rose Li and Associates, Inc.