Measuring Stress, Burnout, and Engagement

There are days that I feel already tired before I go to work.
—Item from the Oldenburg Burnout Inventory,
intended to measure exhaustion

Assessing stress, burnout, and engagement is not as easy as it may seem at first. Think about what you have read in this book so far. How would *you* assess stress? Simple mechanisms for measuring these concepts are not obvious. The question, "Are you stressed?" is an ineffective gauge. People may not know what it means to be stressed, or they may have their own idiosyncratic definition. As you learned in previous chapters, everybody experiences stress differently.

Good data are needed to move forward with intervention efforts. If we don't know how much stress people are experiencing or, more important, their sources of stress, we will have a difficult time reducing it. Moreover, if we want to evaluate our intervention later, we will need a consistent way to assess stress over time. This chapter introduces a variety of options for assessing stress and burnout, focusing on some of the more popular, feasible techniques for healthcare professionals.

ASSESSING STRESS

Historically, a checklist of common dramatic life events (e.g., marriage, divorce, new job, death in the family), called the Social Readjustment Scale, was most commonly used to assess stress (Holmes and Rahe 1967). Each event was given a numeric score (e.g., death of a spouse is 100; divorce is 73; vacation is 13). The higher a person's score, the more stress they were experiencing. While this approach is intuitive, its problem is twofold. First, stress theories argue that stress is experienced within a context. For example, a death in the family could be a welcome end to a long, difficult struggle with cancer and relieve someone from significant caretaking responsibilities. Likewise, in some cases, the scale didn't differentiate between voluntary and involuntary stressors (e.g., initiating a divorce vs. receiving notice that one's spouse filed for divorce). For these reasons, someone can take on a new job, get married, or experience any number of other things and not seem as stressed as someone else because those events may have brought about welcome improvements to that person's previous plight.

Second, the scale was based on major life events. In most cases, pervasive, everyday hassles—as opposed to momentous occurrences—are more commonly the types of stressors that lead to burnout. Perhaps the infusion pump isn't working correctly, or overtime has been made involuntary for the week because of a staff shortage. These types of issues were not reflected in the life events scale, yet they have a profound impact on employees' experiences of stress and strain. In short, the scale lacked sensitivity; someone could have been experiencing high levels of stress at work, but because they did not check off any of the major events listed, they registered low on the stress scale.

In seeking more "objective" approaches to assessing stress, some have advocated for the use of tests that take physiological measures ranging from heart rate to blood cortisol levels. These approaches are somewhat valuable, and there is a host of evidence to support their role in the stress process, but they also have drawbacks. Until recently, physiological measures of stress were difficult to collect,

often requiring blood draws that would not be feasible for use within organizations. That may change as research continues to grow rapidly, validating the use of "wearables" to detect stress primarily through measurement of heart rate variability. However, a recent systematic study from a team of Australian researchers suggests we are still years away from being confident in the reliability of off-the-shelf smart devices for this purpose (Hickey et al. 2021).

Once we get to that point, it is still unclear how helpful those sensors will be. In most cases, the goal is to understand and address *why* employees have these physiological responses, not the responses themselves. Knowing that an employee has high cortisol levels, or a specific heart rate variability pattern, may be a marker that they are stressed, but it does not tell us much about what caused those markers. As a result, we are largely right back at the start in terms of designing any kind of intervention.

Various measures have been created to assess the stressors relevant to healthcare professionals. The Nursing Stress Scale (NSS) has been predominant in the nursing literature (Gray-Toft and Anderson 1981). Widely used in research, this scale has been translated from English into a number of languages, including Spanish and Chinese. In essence, the NSS is a nursing-specific variant of the life event scale that measures stress according to the frequency of stressful events. Since its inception, it has been expanded to capture nine sources of stress in the nursing profession, along with the frequency of each source (see French et al. 2000). This scale has been commonly criticized for neglecting to take into account the resources available to the respondent to address the stressors, similar to the concerns with life events checklists. The applicability of the NSS to healthcare occupations beyond nursing is also questionable.

An alternative approach is to capture the respondent's overall feeling of stress. Instead of focusing on resources, this approach builds the appraisal of demands into the assessment questions. Cohen and colleagues' (1983) Perceived Stress Scale (PSS) is one example of a tool that takes this approach. The PSS and other similar scales are more useful in that they are general enough to apply to most

occupations and tend to return highly valid results. The general nature of these scales' assessment questions, however, makes diagnosis of specific stressors difficult. For example, consider this question: "In the last month, how often have you felt that you were effectively coping with important changes occurring in your life?" Would responses to such a question help pinpoint what needs to be "fixed" to reduce stress in the workplace? Not likely.

The bottom line is that an ideal measure of stress has yet to be designed. The key is to understand whom you are assessing and what you are seeking to measure. If you are trying to get an overall feel for the level of stress in your facility, the PSS may help you gather the information you need to do so. If you want to diagnose specific stressors among professionals of a specific occupation, then a scale similar to the NSS would work better. No matter what you choose, keep in mind that there is no perfect tool for measuring stress among healthcare professionals.

ASSESSING BURNOUT

Although somewhat easier to measure than stress, burnout is similarly difficult to assess. Historically, it has been measured using pencil-and-paper surveys. Following are descriptions of some of the options available including how they approach burnout and their advantages and disadvantages. Many measures have been developed, but their use may be questionable if there is no evidence supporting their effectiveness. For brevity, the discussion covers methods that have produced reliable and valid results. Their outcomes are consistent over time and across items, and they measure what they say they measure.

The Maslach Burnout Inventory

The Maslach Burnout Inventory (MBI), created by Maslach and Jackson (1981), is by far the most-used measure. It is based on

Maslach's three-dimensional conceptualization of burnout, which includes subscales for emotional exhaustion, depersonalization (called *cynicism* in the most recent iteration), and reduced personal efficacy. This measure has been translated into dozens of languages. It is available in several versions including a traditional version for service providers, a general version that does not refer to specific clients and can be applied to nearly any profession, and a version for educators. In recent years, more specific variations of the original measures have been introduced that are better suited for healthcare professionals and students.

The MBI asks the respondent to indicate the frequency with which a series of statements apply to them. Because of its popularity, a great deal of attention has been paid to testing the MBI's reliability and validity. It withstands these tests well; the three subscales emerge as distinct factors. Moreover, it adequately discriminates between burnout and other related measures such as depression.

Despite its common use, the MBI has drawn criticism. The usefulness of its personal efficacy scale is questionable, as it tends not to work as consistently as the other two dimensions. Because it is a commercial test, the MBI involves a cost to administer; organizations have to purchase a license to from Mind Garden, the test's publisher (see www.mindgarden.com). In a large health system, this expense could add up quickly, though there are volume discounts, and the cost is lower if one is using the MBI for research purposes.

The measure's construction poses another problem. All of the items are phrased in the same direction (the respondent's burnout score increases each time they answer a statement affirmatively). People may figure out this pattern and respond without really thinking about the statements. For example, when faced with long surveys, people may repeatedly select the same response (e.g., always select "1"), regardless of how they really feel. Or they may read the first few questions, realize they are all about the same topic, and respond similarly to all questions on the inventory without really thinking about them.

The Oldenburg Burnout Inventory

To address the concerns about cost and structure, and recognizing the limitations of the personal efficacy scale, Demerouti and colleagues (2002) developed the Oldenburg Burnout Inventory (OLBI), which assesses burnout on the basis of two dimensions: emotional exhaustion and disengagement. (The quote at the start of the chapter is a sample item from the exhaustion subscale of the OLBI.) It is free to use for noncommercial purposes. It also addresses the MBI's structure issue by balancing positively and negatively worded items so respondents cannot lapse into an answer pattern but must carefully consider each statement. Originally developed in German, the OLBI has been translated into many languages, validated in peer-reviewed journals, and applies to any occupation; it does not specifically refer to working with customers or patients.

Structured similarly to the MBI, the OLBI includes 16 statements that require responses. Rather than emphasizing the frequency with which the respondent experiences each item, the OLBI asks respondents to indicate the extent to which they agree with the statement. This type of scaling (*strongly agree* to *strongly disagree*) corresponds with many other measures, so it can be consistent when embedded with other measures in a larger survey.

The reliability and validity of the OLBI have also been studied, and it has held up well. A number of studies have compared the OLBI directly to the MBI, finding sufficient conceptual overlap between the emotional exhaustion and depersonalization/ disengagement/cynicism dimensions of the OLBI and MBI.

Other Measures of Burnout

The MBI and OBI are by far the most common measures used to assess burnout in the research literature and in the field. However, I'll briefly note two others.

The Copenhagen Burnout Inventory, developed by Kristensen and colleagues (2005), expands the domain of burnout beyond work to include personal burnout, work-related burnout, and client-related burnout. The idea is to customize the measure to fit the situation of the respondents who are completing the items. For example, if they don't work with clients, there is no need to measure that subsection.

The Staff Burnout Scale for Health Professionals (SBS-HP), as its name suggests, is intended specifically for the healthcare sector (Jones 1980). The SBS-HP measures burnout on the basis of adverse cognitive, affective, behavioral, and psychophysical reactions. Although it has the potential advantage of being tailored to healthcare settings, the SBS-HP does not have the same track record as the measures discussed earlier and its conceptualization of burnout centered on four adverse reactions is inconsistent with much of the literature on the topic.

ASSESSING ENGAGEMENT

The measurement of engagement shares several characteristics with the common measures of stress and burnout. While some options employ formats like checklists to assess conditions underlying engagement, others attempt to measure the psychological processes of engagement. The approach you take with the assessment of engagement depends somewhat on what you are trying to accomplish. As noted in chapter 3, there are two main perspectives on engagement (psychologically focused and management focused); the two most common measures of engagement reflect each of those perspectives.

Gallup Q12

Among the first measures of employee engagement was the Gallup Workplace Audit (Gallup Organization 1992), now called the

Gallup Q^{12}. As its name implies, it purports to cover 12 needs of employees that, when met, will lead to engagement. This measure is similar to the checklist approach to stress; however, rather than focus on whether something has occurred, the Q^{12} emphasizes conditions that would potentially lead to engagement.

One could argue that this measure fits conservation of resources theory very well, as the conditions measured by the Q^{12} are indicators that employees have resources to draw upon to engage with their work. With that in mind, the Q^{12} can be useful if you focus a little less on the extent to which people are engaged in your organization and more on whether the work environment supports engagement.

A drawback of the Q^{12} is that it is a commercial product of the Gallup Organization. The costs can be rather high, especially for a small group. In 2022, the cost of the Q^{12} was $15 per participant, with volume discounts for more than 100 participants. The cost does include additional resources that may be of value, such as benchmarking with other organizations. Like other commercial measures, Gallup claims its measure is reliable and valid. Peer reviews to support those claims are a bit harder to come by, though some data are available (e.g., Harter et al. 2010).

Utrecht Work Engagement Scale

By far, the most common measure of engagement in the academic literature is the Utrecht Work Engagement Scale (UWES) created by Wilmar Schaufeli and colleagues (2002). The UWES measures the psychological experience of engagement, including subscales for vigor, dedication, and absorption facets (discussed in chapter 3). This 17-item scale has been translated into many languages and validated in a large number of peer-reviewed publications. Researchers have adapted the scale for special populations such as students (e.g., Carmona-Halty, Schaufeli, and Salanova 2019).

The UWES is probably a fine option if money is an issue, as it is available without cost. Also, the UWES is useful if you are more interested in measuring the degree of engagement of your employees

than in measuring the conditions that can lead to engagement. However, the UWES may be less useful when it comes time to translate your data into interventions. The items refer more to states of being. For example, if a work unit scored an average of 2 out of 5 on the item "At my work, I feel that I am bursting with energy," what would you do next?

Other Measures of Engagement

Many consulting firms have created their own proprietary engagement measures. Further, other measures have been described in peer-reviewed articles that purport to assess engagement. The differences in these scales reflect differing perspectives on the concept of engagement. For example, the 12-item Shirom-Melamed Vigor Measure (Shirom 2008) concentrates on the vigor facet of engagement, though it divides vigor into physical, emotional, and cognitive components. Bruce Rich and colleagues (2010) came up with a measure for the physical, emotional, and cognitive components of engagement. A measure developed by Thomas Britt and colleagues (2005) includes four items that emphasize performance and commitment elements of the job.

My fellow academics generally justify their creation of new measures of engagement with the assertion that others do not fully measure engagement, even while using the same foundational views of what engagement is. I have gone into less detail on these alternative measures largely because they are used much less frequently in the academic literature and, from what I can tell, in the field.

RECOMMENDATIONS ON THE LENGTH OF ASSESSMENTS

One concern that frequently comes up when I'm working with an organization to collect data concerns the length of the survey.

Many burnout and engagement measures require respondents to complete 15 to 30 items for each component of burnout or engagement (e.g., emotional exhaustion). When considering length of surveys, researchers frequently follow a rule of thumb of 4 items per minute. Typically, people don't use that much time, but that rule captures the vast majority of people completing your survey. If all you are measuring is burnout, or maybe burnout and engagement, you might be fine with the scales as they were designed. However, in many cases we want to capture other information, which will require a longer survey. Particularly if we are concerned that burnout may be a problem in our organization, using a long survey may not be optimal.

Fortunately, you have options. In some cases, researchers have created and tested the validity and reliability of shortened versions of the scale. For example, while the full version of the UWES is 17 items, a 9-item version is available with adequate validity and reliability data to support its use (see Schaufeli, Bakker, and Salanova 2006). More recently, researchers at the University of Alabama tested and found support for the use of a three-item version of the UWES (Matthews, Mills, and Wise 2020) that may be especially attractive if you want to capture frequent, short check-ins with employees to identify changes in engagement over time.

Another approach is to measure more specific facets of burnout and engagement. Because, for example, many believe that exhaustion is the first facet of burnout to manifest, leading to disengagement and reduced personal efficacy, researchers typically suggest measuring only exhaustion. This is the approach I have frequently used in my own work, using either the five-item exhaustion subscale of the MBI or the eight-item exhaustion subscale of the OLBI. As the authors of the MBI note, however, if you are more interested in assessing the point where you might truly consider someone "burned out," it might make more sense to only capture the depersonalization/disengagement/cynicism subscales. Taking this approach is a little less clear for engagement because the three facets of engagement are not temporally structured in quite the same way as burnout.

However, I have seen people concentrate on one or two facets, particularly if there is interest in specific facets.

TAKEAWAY POINTS: GATHERING THE DATA YOU NEED

Although we have some good options for assessing stress and burnout, we don't have one easy, perfect solution. Here are important points to take away from this chapter:

- Stress is exceedingly difficult to measure. There are various imperfect ways to measure stress, and the best tool for the job will depend on the context in which the stress is experienced.
- Measures of burnout are somewhat more consistent than measures of stress. There are important advantages and disadvantages to each.
- Measures of engagement largely take on two forms that are similar to measures of stress and burnout, respectively. They are (1) measures of conditions that should lead to engagement and (2) measures that assess the psychological state of engagement.
- To select an appropriate measure for your situation, you may need to consider shortened versions of the scales or determine which dimensions of construct you wish to assess (e.g., only measuring emotional exhaustion for burnout).

Well, you now have all the background information on stress, burnout, and engagement. Chapter 5 explores solutions to the stress epidemic in healthcare by addressing existing burnout and discussing prevention of additional burnout.