Changes in Sabbath-Keeping and Mental Health Over Time: Evaluation Findings From the Sabbath Living Study

Rae Jean Proeschold-Bell, Beth Stringfield, Jia Yao, Jessica Choi and David Eagle
Duke Global Health Institute, USA

Celia F. Hybels
Duke University Medical Center, USA

Heather Parnell, Kelly Keefe and Sara Shilling
Duke Global Health Institute, USA

Abstract
Work-related stress is experienced at a high level in the United States. Clergy are particularly likely to over-extend themselves to act on their sacred call. Sabbath-keeping may offer a practice that is beneficial for mental health, yet many Protestant clergy do not keep a regular Sabbath. We examined whether United Methodist clergy who attended informative Sabbath-keeping workshops reported changes in spiritual well-being and mental health post-workshop. Compared to baseline, at 3 and 9 months post-workshop, participants reported an increase in Sabbath-keeping. In adjusted random effects and Poisson models, compared to not changing Sabbath-keeping frequency, increasing Sabbath-keeping was related to only one outcome: greater feelings of personal accomplishment at work. Decreasing Sabbath-keeping was related to worse anxiety symptoms, lower spiritual well-being in ministry scores, and a higher probability of having less than flourishing mental health. For four outcomes, there were no significant associations with changes in Sabbath-keeping over time. Although lacking a control group, this study adds to cross-sectional Sabbath-keeping studies by correlating changes in Sabbath-keeping with changes in mental health outcomes over time.

Keywords
intervention, positive psychology, mental health, spiritual well-being, mental illness, Sabbath

Corresponding author:
Rae Jean Proeschold-Bell, Center for Health Policy and Inequalities Research, Duke Global Health Institute, Durham, NC 27708, USA.
Email: rae.jean@duke.edu
Introduction

Stress, defined as a reaction to a stressor and an imbalance between demands and one’s personal and social resources (Lazarus & Folkman, 1984), is a challenge for many. In a survey of a representative sample of Americans ages 16–64, 57% of participants reported being affected by stress, and almost 33% reported seeking out medical care for stress-related concerns (Everyday Health, 2019). The multiple impacts of COVID-19 have only contributed to an increase in stress (Centers for Disease Control and Prevention, 2020).

Although too little stress may prevent one’s stress response system from being robust and stable, prolonged stress can deteriorate the body and brain (McEwen & Gianaros, 2011). Also, high-stress levels are correlated with mental health problems (Toussaint et al., 2016), including depression (Yang et al., 2015) and anxiety (Juruena et al., 2020). With the harm that stress can cause and the increased stress levels nationally, it is critically important to identify ways to mitigate high levels of prolonged stress.

A variety of stressors present before COVID-19 could be contributing to the rise in stress in the United States. In terms of occupational stress, employees feel the effects of our digital age; in particular, demands for attention and multitasking have increased (Moisala et al., 2016; Ziegler et al., 2019). In addition, employees are often expected to be reachable in the evenings and weekends in ways not previously possible. Although a standard work week is thought of as 40 working hours, in the United States, employees often work well above 55 hours weekly (Weston et al., 2019). In one study of a 2-month period, white-collar daytime office workers worked an average of 48 hours per week (Sato et al., 2020). It is possible that stress levels in the United States are rising because many people never fully turn away from work.

One particularly interesting population in terms of stress is clergy. Many clergy experience their occupation as stressful (Chirico, 2017), and stress may be fueling mental health issues for clergy (Ferguson et al., 2015; Proeschold-Bell, Swift, et al., 2013). The potential causes of stress for clergy include clergy’s perception that their endeavors will be diminished if congregants experience them as setting time limits, showing their vulnerability to congregants, or expressing appropriate frustration with congregants (Weaver et al., 2002). In addition, many clergy are observed to work long hours to provide the assistance that their congregants seek (Carroll, 2006). Their devotion to serving could be because clergy experience a sacred calling to their work, such that they consider it to be a vocation rather than an occupation (Campbell, 1994). This calling of clergy is consistent with sanctification theory, which proposes that when something is sacred, people will fiercely protect it, exert extraordinary effort for it, and experience desolation if it is lost (Pargament & Mahoney, 2005). The sacred nature of their work may encourage clergy to over-extend themselves routinely and discourage them from activities that not only mitigate stress but also require time, such as exercise and preparing healthy meals (Proeschold-Bell et al., 2020). In addition to perceived stress among clergy, studies indicate that clergy experience high rates of depression and anxiety (Hybels et al., 2018; Knox et al., 2005; Lau, 2018; Proeschold-Bell, Miles, et al., 2013).

The COVID-19 pandemic may also affect clergy stress. A qualitative study conducted in the United States in June through August 2020 found that the routine ways of conducting ministry were disrupted, leading to clergy modifying their practices during a period of creativity that also felt unsettled (Johnston et al., 2021). Clergy reported difficulties finding ways to conduct pastoral care from a distance and some felt a sense of loss. In a survey study of clergy in England conducted from April through July 2020, clergy were asked about their perceived experience currently versus before the pandemic (Village & Francis, 2021). In total, 37% of clergy indicated experiencing more stress currently whereas 22% indicated less stress; 48% indicated more exhaustion whereas 19% indicated less exhaustion; and 46% of clergy indicated more frustration whereas 11% indicated experiencing less frustration currently. At the same time, high percentages of clergy reported feeling more thankful, more creative, and more prayerful currently.
Among the possible solutions to decrease the effects of stress are various approaches to stress management, including exercise (Puterman et al., 2018; Stubbs et al., 2017; Stults-Kolehmainen & Sinha, 2014), meditation (Dunlop, 2015; Househam et al., 2017), expressive writing (De Moor et al., 2008), diaphragmatic breathing (Brown et al., 2013; Hopper et al., 2018; Perciavalle et al., 2017), music therapy (Umbrello et al., 2019), and guided imagery (Giacobbi et al., 2018). Generally, these approaches consist of small practices occurring multiple times per week. It is possible that a different approach, namely a way of structuring one’s time that minimizes work intrusions for an extended period of time, would be of similar or greater benefit to reduce stress and accompanying mental health problems.

With our interest in clergy, specifically Protestant Christian clergy, we wondered if the spiritual practice of Sabbath-keeping would benefit clergy well-being. Sabbath-keeping, rooted in the story of Creation, is a Judeo-Christian tradition that mirrors the account that God rested on the seventh day after creating the world in 6 days (Speedling, 2019). While Orthodox Jews refrain from all work on the Sabbath, in contemporary Protestant circles, Sabbath is often defined as a day to pray and play (Peterson, 2018), which may include resting, spending time in nature, and disconnecting from work and unenjoyable tasks and, instead, connecting with loved ones and God. For some, Sabbath is a dedicated time to reflect upon their daily lives (Woodley, 2015). White et al. (2015) propose that truly recognizing the Sabbath means engaging in “virtuous rest,” which orients hearts and minds toward God. However, the degree to which a person observes the Sabbath can vary widely.

In terms of mental health, Sabbath-keeping has the potential to guide people to focus on what is most important in life (Dein & Loewenthal, 2013), temporarily opt-out of an otherwise pervasive “anxiety system” (Brueggemann, 2014), and enable individuals to reflect on the actions they take in a complicated, changing world (Muller, 2000). Sabbath-keeping can foster gratitude (Carter, 2013; Dein & Loewenthal, 2013). Greater Sabbath-keeping has been associated with mental health benefits, including reduced depressive symptoms (Pargament, 2011), better self-control (Dein & Loewenthal, 2013), and promotion of a sense of personal empowerment (Goldberg, 1986). A mediational analysis also revealed that Sabbath-keeping correlated strongly with higher mental health functioning (Superville et al., 2014).

Although the few empirical studies noted earlier point to potential mental health benefits of Sabbath-keeping, the studies leave several gaps, including an ambiguous definition of a Sabbath-keeper (White et al., 2015) and the inability to interpret the temporal associations between Sabbath-keeping and other variables (Superville et al., 2014). In addition, these studies were not done with clergy and the potential for mental health benefits for clergy may be different. Furthermore, we did not find any experimental studies on the relationship between Sabbath-keeping practices and spiritual well-being, despite the fact that people who choose to keep Sabbath likely value spiritual well-being.

In this study, we build upon the available literature by investigating the correlations between Sabbath-keeping, mental health, and spiritual well-being over time. We also explore these relationships among clergy, whose vocation remains steeped in religion, although not necessarily with time easily carved out for religious practices like Sabbath-keeping. Our research team served as the external evaluators for a non-governmental organization (NGO) that provides didactic Sabbath-keeping workshops to clergy. In this exploratory study, we collected evaluation data to examine clergy participant changes in spiritual well-being, positive mental health, and mental health problems across 9 months.

**Methods**

**Recruitment**

Participant inclusion criteria were being United Methodist Church (UMC) clergy with a current or previous appointment in the North Carolina or Western North Carolina Annual Conference, the two
geographic regions into which UMC clergy in the state are organized. Thus, participants included elders, local pastors, deacons, and retired clergy. Student pastors (pastors who are currently in seminary but are the lead or associate pastor for a church for 20 hours/week) with a UMC appointment were also eligible.

Our NGO partner held workshop events from February 2017 to November 2018. Event inclusion criteria were the following: minimum of 3 hours in duration, held in North Carolina, conducted by the staff of our NGO partner, and attended primarily by clergy. Our NGO partner recruited for the intervention through multiple means, including clergy email lists, conference-level and district-level communications such as e-newsletters, and in-person networking and word-of-mouth. In addition, many of the workshop events were endorsed by district superintendents (i.e., clergy who supervise the 60–120 clergy serving in a particular geographic area) and some workshops were offered as district events for which the district superintendent would expect all clergy in that region to attend. Sometimes, the workshop event was held adjacent to other clergy programming, such as the conference-wide Annual Conference meeting. Clergy could attend more than one workshop, although this occurred in less than 5% of cases. For evaluation purposes, when a participant attended more than one workshop and completed the evaluation surveys more than once, we used their more complete survey data set. If more than one set of surveys was fully completed by an individual, we used the first completed survey set.

**Sabbath-keeping intervention (“workshop events”)**

The NGO Blessed Earth created and delivered workshop events of varying lengths, from 3 to 15 hours. Workshop content focused on the reasons to keep Sabbath, and included Bible passages, prayer, anecdotes of how keeping Sabbath has benefited individuals, couples, and families, and envisioning how regular Sabbath-keeping may benefit congregants and the church as a whole. Content varied and was tailored to the event and audience. Core content emphasized the busyness and potential harm of today’s current lives that are non-stop and overly full, and Sabbath-keeping was held up as a healing anecdote offered by God and designed to be observed in community with each other. Events included some discussion of common practical questions around what activities might and might not fit in a Sabbath day. A small amount of time was devoted to planning how to enact Sabbath-keeping personally. Participants were given resources, including books, videos, Sabbath plan worksheets, and Sabbath scriptures and quotations.

**Data collection**

Clergy registered for the workshop and then our NGO partner provided the research team with clergy names and email addresses. The research team emailed them an invitation to participate in the study and a link to an online consent form; they were prompted multiple times in advance of the workshop to participate in the study. Following consent, participants completed an online survey. Clergy who attended the workshop but who had not consented and completed the survey were recruited by NGO partner staff who had a verbal script and a flyer with the survey link. Workshop attendees who met inclusion criteria could enroll in the study if they consented and completed the baseline survey within 7 days of the workshop. In a few cases, surveys were completed on paper. The self-administered survey took 20–40 minutes to complete and consisted of a variety of mental health and Sabbath-keeping measures. Participants who completed at least half of the survey within a week of their workshop were asked to take the survey again 3 and 9 months later. Participants were not compensated for time spent taking their baseline survey, but they received Amazon gift cards in the amounts of US$25 for the 3-month survey and US$35 for the 9-month survey.
All procedures were approved by the Duke University Campus Institutional Review Board and all participants gave informed consent.

**Measures**

*Sabbath-keeping* was measured by the item, “In the past 4 weeks, how many times did you keep Sabbath?” Sabbath-keeping was defined as, “By Sabbath, we mean anything you considered to be Sabbath, no matter how you spent your time and even if it was less than 24 hours.”

*Spiritual well-being* was measured using the Clergy Spiritual Well-being Scale, which has separate scales for spiritual well-being in daily life (seven items, for example, “How often do you experience the presence and power of God in: the ordinary?; your closest relationships?; your personal prayer life?”), and spiritual well-being in ministry (nine items, for example, “How often do you experience the presence and power of God when sharing the sacraments?; leading worship?; conducting pastoral visitations; sharing in crisis intervention?”; Proeschold-Bell, Miles, et al., 2013). Mean scores for each scale could range from 0 to 4. Spiritual well-being in ministry items were only asked of clergy serving congregations, resulting in a slightly smaller sample size for that measure.

*Work-related burnout* was measured using the Maslach Burnout Inventory (MBI), Human Services version (Maslach & Jackson, 1981). The three components of burnout conceptualized by Maslach (1993) and included as separate scales are emotional exhaustion (nine items), cynicism toward or depersonalization of the people one serves (five items), and personal accomplishment (eight items). Mean scores could range from 0 to 6.

*Anxiety* was measured using the Generalized Anxiety Disorder-7 (GAD-7, Spitzer et al., 2006), which consists of seven items asking about the past 2-week frequency of experiencing anxiety symptoms, such as feeling nervous. Mean scores could range from 0 to 3.

*Quality of one’s spousal or romantic relationship* was measured using the Quality of Relationship Index (Norton, 1983). The six items ask about one’s overall satisfaction with their relationship. We changed the original wording from “marriage” to “relationship” to include couples that are not married. Mean scores could range from 0 to 4, with higher scores indicating more relationship satisfaction.

*Positive mental health* was measured using the Mental Health Continuum-Short Form (MHC-SF), which is a 14-item scale asking about the frequency of experiences in the past month (Keyes, 2002). Three items capture emotional well-being (e.g., “happy”), six items capture psychological well-being (e.g., “you had warm and trusting relationships with others”), and five items capture social well-being (e.g., “you had something important to contribute to society”). We created dichotomous scores of flourishing (high positive mental health) and less than flourishing (languishing plus moderate mental health) based on instructions by Keyes (2011).

*Demographic items* included gender, years in ministry (“How many years have you been in ministry full or part-time for which you were paid a salary?”), bi-vocational status (“Do you consider yourself to be bi-vocational?”), children currently living at home (“Do you have children of any age living at home?”), and number of congregations appointed to (“How many congregations are included in your current appointment?”).

**Statistical analysis**

Participants who reported their Sabbath-keeping frequencies at baseline and at least one follow-up survey were included for analysis. We examined the longitudinal data and calculated summary statistics, namely means and standard deviations for continuous variables, and counts and percentages for binary and categorical variables. For demographic and occupational variables, we reported
the summary statistics at baseline; for Sabbath-keeping frequencies and the well-being outcomes, we reported the summary statistics over time.

For each continuous well-being outcome, a random effects model was fit and it estimated associations of the outcome at follow-up with the covariates. The continuous well-being outcomes included spiritual well-being in the everyday, spiritual well-being in ministry, anxiety symptom severity, quality of relationship with one’s romantic partner, emotional exhaustion, depersonalization, and personal accomplishment. For the binary well-being outcome, flourishing, a modified Poisson model was fit and allowed us to report prevalence ratios (Zou, 2004).

Covariates in each model included the outcome of interest at baseline, Sabbath-keeping frequency at baseline, change in Sabbath-keeping frequency from baseline to follow-up, follow-up time point, and interaction between change in Sabbath-keeping frequency from baseline to follow-up and follow-up time point, as well as the control variables. Change in Sabbath-keeping frequencies were modeled as categorical variables, in which we estimated associations with Sabbath-keeping both for increasing and decreasing the number of times one kept Sabbath in the past 4 weeks, compared to participants that kept Sabbath at constant frequencies over time.

The control variables in the final models included gender, bi-vocational career status (i.e., concurrently working in ministry and having another job), and experience in ministry. Age and experience in ministry (measured in number of years) are highly correlated and therefore could not both be included in the model. We included experience in ministry because it is conceptually more relevant to Sabbath-keeping than age.

Two additional control variables, whether the participant had any children living at home and the number of congregations the participant was appointed to, were included in earlier model specifications and then excluded from the final models, because these variables were conceptually less relevant and were not found to be statistically significantly associated with any of the outcomes. The relationships of main analysis interest (i.e., associations of outcomes with change in Sabbath-keeping frequency from baseline to follow-up, follow-up time point, and interaction between change in Sabbath-keeping frequency from baseline to follow-up and follow-up time point) were similar in models including and excluding children living at home and number of congregations.

For each outcome, a model specification including an additional square term of Sabbath-keeping frequency at baseline was fit. If the outcome was significantly associated with the square term, it was included in the final model specification. Each model specified robust standard errors to account for heteroskedasticity. Finally, we calculated the effect size for the association of each continuous outcome with change in Sabbath-keeping frequency from baseline to follow-up by standardizing the outcome by time point then re-running the model, yielding the standardized coefficient for the association.

**Results**

The program team conducted 17 Sabbath-keeping events from which clergy were recruited for the study. Six of the events included 3–4 hours of content; eight of the events included more hours of content and an overnight experience.

A total of 227 clergy consented to the study. Twenty-two participants were excluded for not answering Sabbath-keeping survey questions at baseline, and an additional 38 participants were excluded for not answering Sabbath-keeping survey questions at either of the follow-up time points, leaving 167 clergy who reported on their Sabbath-keeping both at baseline and at least one follow-up time point. Of those 167 participants, 160 took the 3-month follow-up survey and 136 took the 9-month follow-up survey. In total, 129 participants answered the Sabbath-keeping questions in all three surveys. The majority of participants were female (60%, see Table 1), which is high compared to the population of UMC clergy in North Carolina (34% female). Participants reported a range of years of experience in vocational ministry, with 31% reporting 10–19 years of
experience. In total, 7% were retired (11 participants). Nearly one-fifth (17%) indicated they were bi-vocational, meaning they concurrently held a job in addition to doing paid ministry work, compared to 12% of UMC clergy in North Carolina.

Before testing for a significant association between changes in Sabbath-keeping frequency and changes in mental health outcomes, we examined whether participants changed their Sabbath-keeping over time, and also whether outcomes changed over time. Of the 167 participants who reported their recent Sabbath-keeping behavior on the baseline survey, 25.7% indicated keeping Sabbath four times over 4 weeks. This percentage was higher at 9 months (45.6%, see Figure 1). The percentage of participants who reporting keeping no weeks of Sabbath in the past 4 weeks at baseline was 21.0%, which decreased to only 8.8% at 9 months. Thus, in aggregate, the frequency of Sabbath-keeping in

Table 1. Description of Participants at Baseline (N = 167).

<table>
<thead>
<tr>
<th></th>
<th>%</th>
<th>n</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Gender</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>40.1</td>
<td>67</td>
</tr>
<tr>
<td>Female</td>
<td>59.9</td>
<td>100</td>
</tr>
<tr>
<td><strong>Any children living at home</strong></td>
<td>40.7</td>
<td>68</td>
</tr>
<tr>
<td><strong>Bi-vocational</strong></td>
<td>16.8</td>
<td>28</td>
</tr>
<tr>
<td><strong>Experience in vocational ministry</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0–4 years</td>
<td>24.0</td>
<td>40</td>
</tr>
<tr>
<td>5–9 years</td>
<td>13.2</td>
<td>22</td>
</tr>
<tr>
<td>10–19 years</td>
<td>31.1</td>
<td>52</td>
</tr>
<tr>
<td>20–29 years</td>
<td>15.6</td>
<td>26</td>
</tr>
<tr>
<td>30 years and above</td>
<td>16.2</td>
<td>27</td>
</tr>
<tr>
<td><strong>Number of congregations appointed to</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>None</td>
<td>15.0</td>
<td>25</td>
</tr>
<tr>
<td>1 congregation</td>
<td>61.7</td>
<td>103</td>
</tr>
<tr>
<td>2 or more</td>
<td>23.4</td>
<td>39</td>
</tr>
</tbody>
</table>

Figure 1. Sabbath-keeping at baseline, 3 months, and 9 months.
the past 4 weeks increased. However, at an individual level, it was possible for participants to decrease the number of times they kept Sabbath in the past 4 weeks.

The unadjusted, aggregate scores of several mental health variables indicated four statistically significant improvements between baseline and 9 months (see Table 2). Mean scores of spiritual well-being in the everyday increased from 2.72 (SD = 0.81) at baseline to 2.82 (SD = 0.74) at 9 months. Mean scores of spiritual well-being in ministry increased from 2.76 (SD = 0.76) at baseline to 2.94 (SD = 0.69) at 9 months. One kind of burnout, emotional exhaustion, had scores that improved slightly at 3 months (1.85, SD = 1.27 at baseline vs. 1.71, SD = 1.23 at 3 months) but returned to baseline levels at 9 months (1.86, SD = 1.24). Overall, changes were small.

We tested whether changes in the frequency of Sabbath-keeping over time were significantly associated with changes in multiple mental health outcomes over time, adjusting for gender, being bi-vocational, and number of years of ministry experience (see Table 3). None of the adjusted variables were significantly associated with any of the outcomes.

Of the significant relationships between change in well-being outcomes and change in Sabbath-keeping, decreasing the number of times one kept Sabbath in the past 4 weeks related to worse mental health. Specifically, decreasing Sabbath-keeping by 2–4 times between baseline and 3 months, compared to not changing Sabbath-keeping frequency, significantly related to worse anxiety scores ($B = 0.411, 95\% \text{ CI} = [0.063, 0.760]; p = .021$), with an effect size between medium and large (standardized $\beta = 0.70$; Sullivan & Feinn, 2012). Decreasing Sabbath-keeping by just one time in the past 4 weeks between baseline and 3 months, compared to not changing Sabbath-keeping, also significantly related to worse anxiety scores ($\beta = 0.293, 95\% \text{ CI} = [0.132, 0.454]; p < .001$), with a medium effect size (standardized $\beta = 0.51$). At 9 months, decreasing Sabbath-keeping by 2–4
Table 3. Model of Change in Well-Being Outcomes Over Time With Change in Sabbath-Keeping Over Time.

<table>
<thead>
<tr>
<th>Sample size (number of participants; number of observations)</th>
<th>SWB in ministry</th>
<th>Anxiety symptoms</th>
<th>Personal accomplishment scores</th>
<th>PR [95% CI], p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outcome at baseline</td>
<td>125; 217</td>
<td>164; 290</td>
<td>158; 276</td>
<td>163; 288</td>
</tr>
<tr>
<td>Sabbath-keeping frequency at baseline, centered around two times per 4 weeks</td>
<td>0.776 [0.680, 0.872]; p &lt; .001</td>
<td>0.633 [0.554, 0.711]; p &lt; .001</td>
<td>0.713 [0.572, 0.854]; p &lt; .001</td>
<td>4.544 [2.606, 7.922]; p &lt; .001</td>
</tr>
<tr>
<td>Linear term</td>
<td>0.036 [-0.024, 0.096]; p = .241</td>
<td>-0.037 [-0.087, 0.014]; p = .154</td>
<td>0.085 [-0.006, 0.176]; p = .067</td>
<td>0.880 [0.744, 1.041]; p = .136</td>
</tr>
<tr>
<td>Square term</td>
<td>0</td>
<td>-0.069 [-0.104, -0.034]; p &lt; .001</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Change in SK frequency from baseline to follow up (ref: no change), with reference time point being 9 months follow-up</td>
<td>Decreased for 2–4 times per 4 weeks</td>
<td>-0.528 [-0.994, -0.063]; p = .026</td>
<td>0.042 [-0.256, 0.341]; p = .782</td>
<td>-0.320 [-1.121, 0.481]; p = .433</td>
</tr>
<tr>
<td></td>
<td>Decreased for 1 time per 4 weeks</td>
<td>0.040 [-0.148, 0.227]; p = .680</td>
<td>-0.044 [-0.277, 0.190]; p = .715</td>
<td>-0.350 [-0.799, 0.098]; p = .126</td>
</tr>
<tr>
<td></td>
<td>Increased for 1 time per 4 weeks</td>
<td>-0.072 [-0.274, 0.130]; p = .486</td>
<td>-0.087 [-0.278, 0.105]; p = .375</td>
<td>-0.064 [-0.262, 0.135]; p = .531</td>
</tr>
<tr>
<td></td>
<td>Increased for 2–4 times per 4 weeks</td>
<td>0.182 [-0.049, 0.413]; p = .122</td>
<td>-0.076 [-0.310, 0.157]; p = .520</td>
<td>0.097 [-0.214, 0.408]; p = .542</td>
</tr>
<tr>
<td>Change in SK frequency from baseline to follow up (ref: no change), from re-running the models with reference time point being 3 months follow-up</td>
<td>Decreased for 2–4 times per 4 weeks</td>
<td>-0.287 [-0.723, 0.150]; p = .186</td>
<td>0.110 [-0.063, 0.370]; p = .219</td>
<td>-0.363 [-0.832, 0.107]; p = .130</td>
</tr>
<tr>
<td></td>
<td>Decreased for 1 time per 4 weeks</td>
<td>0.081 [-0.317, 0.468]; p = .689</td>
<td>0.293 [-0.132, 0.645]; p &lt; .001</td>
<td>-0.206 [-0.676, 0.264]; p = .391</td>
</tr>
<tr>
<td></td>
<td>Increased for 1 time per 4 weeks</td>
<td>-0.044 [-0.245, 0.158]; p = .671</td>
<td>-0.025 [-0.196, 0.046]; p = .774</td>
<td>-0.038 [-0.301, 0.225]; p = .778</td>
</tr>
<tr>
<td></td>
<td>Increased for 2–4 times per 4 weeks</td>
<td>0.230 [-0.015, 0.475]; p = .066</td>
<td>-0.021 [-0.204, 0.162]; p = .819</td>
<td>0.277 [0.001, 0.554]; p = .049</td>
</tr>
<tr>
<td>3 months follow-up (ref: 9 months follow-up)</td>
<td>-10.15 [-0.265, 0.553]; p = .200</td>
<td>-0.087 [-0.230, 0.056]; p = .232</td>
<td>-0.056 [-0.242, 0.129]; p = .553</td>
<td>0.871 [0.592, 1.280]; p = .482</td>
</tr>
<tr>
<td>Change in SK frequency × 3 months follow-up</td>
<td>Decreased for 2–4 times per 4 weeks × 3 months follow-up</td>
<td>0.554 [0.077, 1.031]; p = .203</td>
<td>0.369 [-0.093, 0.831]; p = .118</td>
<td>-0.043 [-0.964, 0.879]; p = .928</td>
</tr>
<tr>
<td></td>
<td>Decreased for 1 time per 4 weeks × 3 months follow-up</td>
<td>0.042 [-0.404, 0.488]; p = .854</td>
<td>0.337 [0.061, 0.612]; p = .017</td>
<td>0.144 [-0.504, 0.793]; p = .663</td>
</tr>
<tr>
<td></td>
<td>Increased for 1 time per 4 weeks × 3 months follow-up</td>
<td>0.028 [-0.198, 0.255]; p = .807</td>
<td>0.062 [-0.139, 0.262]; p = .547</td>
<td>0.026 [-0.277, 0.328]; p = .868</td>
</tr>
<tr>
<td></td>
<td>Increased for 2–4 times per 4 weeks × 3 months follow-up</td>
<td>0.048 [-0.241, 0.337]; p = .745</td>
<td>0.055 [-0.180, 0.290]; p = .645</td>
<td>0.181 [-0.112, 0.474]; p = .226</td>
</tr>
<tr>
<td>3 months follow-up (ref: 9 months follow-up)</td>
<td>Male (ref: female)</td>
<td>0.024 [-0.141, 0.190]; p = .771</td>
<td>0.021 [-0.083, 0.124]; p = .693</td>
<td>0.136 [-0.083, 0.355]; p = .222</td>
</tr>
<tr>
<td></td>
<td>Bi-vocational at baseline (ref: single-vocational)</td>
<td>0.018 [-0.190, 0.226]; p = .866</td>
<td>-0.009 [-0.132, 0.115]; p = .892</td>
<td>0.001 [-0.229, 0.232]; p = .991</td>
</tr>
<tr>
<td></td>
<td>Experience in vocational ministry at baseline (in years, centered around 14 years)</td>
<td>-0.007 [-0.014, 0.001]; p = .084</td>
<td>0.001 [-0.004, 0.006]; p = .692</td>
<td>0.000 [-0.009, 0.010]; p = .956</td>
</tr>
</tbody>
</table>

CI: confidence interval; PR: prevalence ratio; SK: Sabbath-keeping; SWB: spiritual well-being.
times between baseline and 9 months, compared to not changing Sabbath-keeping, related to worse spiritual well-being in ministry scores ($\beta = -0.528$, 95% CI = $[-0.994, -0.063]$; $p = .026$), with a large effect size (standardized $\beta = -0.86$) and a higher probability of having less than flourishing mental health (PR = 1.975, 95% CI = 1.137, 3.431; $p = .016$).

Compared to decreasing Sabbath-keeping, we found fewer statistically significant associations between change in outcomes and increasing Sabbath-keeping over time. Feelings of personal accomplishment at work were greater for those who increased their Sabbath-keeping 2–4 times per 4 weeks between baseline and 3 months, compared to not changing Sabbath-keeping frequency ($B = 0.277$, 95% CI = $[0.001, 0.554]$; $p = .049$; effect size = 0.26).

There were no significant associations found between changes in Sabbath-keeping over time and changes in: quality of relationship with one’s romantic partner, spiritual well-being in the everyday, emotional exhaustion, or depersonalization (results not shown).

To see if larger increases in Sabbath-keeping related to outcomes, we examined data from the participants who were naïve to Sabbath-keeping at baseline and had kept Sabbath 3–4 times in the past 4 weeks at either 3 or 9 months. Only 17 participants met these criteria. On just these participants compared with all others, we did an ad hoc analysis from baseline to 3 months and baseline to 9 months. We found that emotional exhaustion significantly improved for these participants who went from no to high practice, compared to everyone else, at 3 and 9 months. We did not find significant differences for six other continuous outcomes, although most of them were in the expected direction. These findings are consistent with an interpretation of a beneficial but small relationship between Sabbath-keeping and mental health outcomes.

**Discussion**

In this longitudinal study of changes in Sabbath-keeping among clergy who attended a Sabbath-keeping workshop event, we observed some significant correlated changes of Sabbath-keeping over time with mental health outcomes at 3 and 9 months. One significant association was found with an increase in Sabbath-keeping frequency, and that was feeling greater accomplishment at work at 3 months. Also, participants naïve to Sabbath-keeping and who greatly increased their Sabbath-keeping experienced less emotional exhaustion over time. More significant associations were found over time among participants who indicated a decrease in Sabbath-keeping frequency. For them, decreases in Sabbath-keeping were associated over time with worse mental health (higher anxiety scores, lower spiritual well-being scores, and lower prevalence of qualifying as having flourishing mental health).

Existing studies of populations other than clergy are mainly cross-sectional and have examined increases in Sabbath-keeping frequency, not decreases. These studies have found positive associations between Sabbath-keeping and outcomes such as spirituality development (Speedling, 2019), improved mental and physical health functioning (Superville et al., 2014), and strengthening relationships with others (Dein & Loewenthal, 2013). We therefore hypothesized that we would find positive relationships in this study. There are many potential explanations for the null findings we found for participants who increased their Sabbath-keeping frequency. These explanations include clergy already lead spiritual lives, such that a Sabbath day may be less impactful for their spiritual well-being over time; the emotional and fragmented nature of clergy work may require a more intensive intervention than a Sabbath day without prescribed activities; and the effort to organize one’s week to preserve a day may offset longitudinal mental health gains. Our findings may differ from previous findings because other studies have been cross-sectional, not of clergy, or focused on different denominations (e.g., Seventh Day Adventists). It is also possible that clergy may acclimate to having a Sabbath day over time, thus perceiving no difference in their mental health symptoms 3 and 9 months later, which may show up as lack of improvement on the mental health survey measures. Consistent with this latter hypothesis, our data indicated that decreases in
Sabbath-keeping frequency were significantly associated over time with a slight worsening of mental health. Potentially, once Sabbath-keeping is removed, clergy may feel the effects of its loss on their mental health. Alternatively, because our study design does not allow for causal inference, it is possible that an external stressful factor caused both a decrease in Sabbath-keeping and a slight worsening of some mental health indicators.

This study is one of the first that looked specifically at changes in Sabbath-keeping and their relationship with changes in spiritual well-being. As noted earlier, we found only one significant relationship and that was at 9 months among clergy who had decreased their Sabbath-keeping 2–4 times in the past 4 weeks. It is possible that clergy have many personal and work-related activities that increase—or decrease—their spiritual well-being, and that the effect of increasing Sabbath-keeping washes out in the face of these other activities. Alternatively, we did not measure what clergy spent their Sabbath day doing, and it is possible that many clergy treated the day as a day away from work and did not engage in spiritual practices that day, or decreased their spiritual practices on other days of the week to preserve a day for Sabbath.

We also learned that clergy were interested in Sabbath-keeping and willing to attend workshops on it. Sabbath-keeping workshops were especially attractive to female clergy, given that approximately 70% of the UMC clergy in North Carolina are male (Bethune, 2017) and yet, 60% of this study’s participants were female. Some speculations on why female clergy were attracted to Sabbath-keeping include the following: (1) Female clergy may perceive a greater need for interventions that involve rest, given the excess hours women spend on work plus family responsibilities compared to men (Gjerdingen et al., 2000); (2) women who enrolled may have recruited their female clergy friends to participate, as networking was encouraged and female clergy may find support from other female clergy; and (3) Sabbath Living was at its core spiritual formation work and fundamentally relational, which may be especially appealing to women. Future researchers should explore the reasons for gender differences in Sabbath-keeping interest, and this may be particularly important given the disproportionately negative impact of the COVID-19 pandemic on working women (Connor et al., 2020).

In addition to attending workshops, participants in aggregate increased their Sabbath-keeping. The percentage of participants who did not keep Sabbath at all decreased, and the percentage who kept Sabbath four times per month increased. Many of the workshops were of short duration (e.g., 3 hours), making these changes in Sabbath-keeping frequency noteworthy.

Although this study focuses on outcomes associated with Sabbath-keeping, the process involved in adopting Sabbath-keeping is complex, and known barriers to weekly Sabbath-keeping for clergy include being bi-vocational and experiencing extremely high time demands (Hough et al., 2019). Keeping Sabbath involves reorganizing all non-restful activities into the other 6 days of the week to preserve a day of rest. It involves choices, communicating with many people, proactive planning, and a certain degree of subjectivity as to what is life-giving and acceptable to do on one’s Sabbath. A number of people have written about this topic and given suggestions (Muller, 2000; Sleeth, 2012), and they note that Sabbath-keeping was intended to be practiced as a community in ways not likely experienced by this study’s participants. If Sabbath could be experienced as a community, the proactive communication and decision-making would be substantially less.

**Strengths and limitations**

The major strength of this study is the longitudinal study design. Previous studies that examined the relationship between Sabbath practices and mental health have been cross-sectional, limiting them from determining a causal relationship or associations over time. This study’s longitudinal design, albeit lacking a control group, allowed us to examine directional changes in Sabbath-keeping and their associations with outcomes over time. Another strength of the study is the
inclusion of several mental health outcomes, including positive outcomes such as positive (flourishing) mental health and spiritual well-being.

This study also has several limitations. First, by using a convenience sampling method to identify the clergy members who were interested in participating in the study, external validity is reduced. Second, the researchers did not define what qualified as Sabbath for the participants, and therefore, the term was self-defined. There was likely a wide range of activities that participants considered to be Sabbath-keeping, diverse in both whether spiritual practices were included and whether non-work tasks that are not restful, like errands, were included. In addition, we measured the number of times per month participants kept Sabbath and not the number of hours. It is possible that the relationship to mental health outcomes would have been stronger if we had been able to account for the diversity of Sabbath-keeping experiences. For example, if someone spent time doing errands and considered it Sabbath-keeping, they would be less likely to experience changes in spiritual well-being compared to someone who spent time engaged in spiritual practices on their Sabbath. The intervention itself also varied in that the number of workshop hours ranged from 3 to 15 but this was not measured by the research team, precluding us from examining findings for participants who received more versus less intervention. Third, the study lacks a control group, making it impossible to infer causality. Even though our findings are correlational over time, we cannot confirm directionality; it is possible that participants experiencing stress decreased their Sabbath-keeping and simultaneously experienced increases in anxiety symptoms that were due to stressful conditions and not changes in Sabbath-keeping. Finally, given our small sample size and exploratory nature of the study, we did not adjust for alpha inflation. Had we adjusted the p value to be significant at .00625 using a Bonferroni correction (.05 divided by eight tests), all but one of the significant findings would no longer be significant, and as such, we have interpreted them conservatively. Experimental studies of Sabbath-keeping with larger sample sizes are needed.

Future research

While our findings do not give a strong signal about the relationship between Sabbath-keeping and mental health, this longitudinal study adds new information to prior cross-sectional work. Even with recruiting from the full population of clergy and not just those in need of mental health improvement, we found some signs of benefit over time. The next step for those interested in studying the effects of Sabbath-keeping would be to recruit a sample of non-Sabbath-keepers who are also experiencing undesirable mental health symptoms and conduct an intervention trial. It would be helpful to use a strict definition of hours of Sabbath-keeping and a clear measurement of what participants do and do not do on their Sabbath days. We recommend data collection at multiple follow-up time points, because the stressor of reorganizing one’s week may decrease over time with experience, and mental health outcomes may be better later on, if Sabbath-keeping is maintained. It would be good to control for participants’ stress, available time per week, and willingness to try new things. Such a trial would allow increased certainty that changes in mental health are due to changes in Sabbath-keeping. For researchers wishing to study Sabbath-keeping in clergy, more research is needed on how clergy manage to keep Sabbath when the Sabbath day indicated in scripture is also a work day for them (Hough et al., 2019).

Conclusion

This study adds to cross-sectional Sabbath-keeping studies by additionally correlating changes in Sabbath-keeping over time with changes in mental health outcomes over time. We observed a few small but statistically significant relationships between changes in Sabbath-keeping behavior and changes in outcomes over time, all in a beneficial direction. Notably, cutting back on one’s Sabbath-keeping appeared as significant for more outcomes than increasing one’s Sabbath-keeping. Further
studies with strong inferential designs would expand our knowledge about the effects of changing Sabbath-keeping on mental health over time.

Acknowledgements

The authors thank the developers and implementers of the Sabbath Living intervention. The focus on Sabbath-keeping was brought forth by Matthew Sleeth and Nancy Sleeth, and translated into an intervention by the team at Blessed Earth, led by Nancy Sleeth and Guy Brewer, with key contributions by Denise Baker, Lisa Bachman, Adam York, Susanne Dornsmith, Cynthia Vaughn, and Ivy Parker. They also thank the Duke Global Health Institute’s Research Design and Analysis Core for statistical consultations and the Clergy Health Initiative Writing Group for manuscript comments.

Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study was funded by a grant awarded to Blessed Earth by the Rural Church Area of The Duke Endowment.

ORCID iDs

Rae Jean Proeschold-Bell https://orcid.org/0000-0003-2008-3053
David Eagle https://orcid.org/0000-0003-4909-9497

References


**Author Biographies**

**Rae Jean Proeschold-Bell**, PhD, is Research Director of the Duke Clergy Health Initiative and Research Professor in the Duke Global Health Institute and the Duke University Center for Health Policy and Inequalities Research. Her research interests are in clergy health, well-being practices, and intervention research.

**Beth Stringfield**, is a Research Program Leader in the Center for Health Policy and Inequalities Research within the Duke Global Health Institute. Her research interests include access to care, health disparities, and community-based research.

**Jia Yao**, MA is an Associate in Research at Duke University Center for Health Policy and Inequalities Research. She works on the Duke Clergy Health Initiative team as a data manager and statistical analyst and enjoys making a difference in the lives of clergy.

**Jessica Choi**, MSc, is a Research Analyst at the Duke Global Health Institute and the Duke University Center for Health Policy and Inequalities Research. Her research interests include health disparity, community health, and mental health.

**David Eagle**, PhD, is assistant research professor in the Global Health Institute at Duke University. His research is motivated by a desire to improve the state of religious clergy in contemporary settings and understand the factors that help congregations thrive. Other research interests include how the changing religious landscape in North America is impacting theological education, congregations, and religious leaders.

**Celia F. Hybels**, PhD, is an epidemiologist and Research Associate Professor Emeritus in the Duke University Department of Psychiatry and Behavioral Sciences as well as a Senior Fellow in the Duke University Center for the Study of Aging and Human Development. Her research interests include the epidemiology of psychiatric disorders in late life, psychosocial correlates of and risk factors for depressive symptoms, and statistical modeling of trajectories of depression.

**Heather Parnell**, MSW, is a Research Project Manager at the Duke University Center for Health Policy and Inequalities Research. She conducts health disparities research using mixed methods.

**Kelly Keefe**, MHA, is a Clinical Trials Project Leader for the Duke Clinical Research Institute. Her research interests are stakeholder engagement and developing robust participant-centric research strategies.

**Sara Shilling**, MSW, MPH, LCSWA, is a clinical social worker with a focus on families and youth. She is interested in direct clinical service, program planning, and evaluation to support the advancement of families and youth.