Unpacking Habit With Bayesian Mixed Models: Dynamic Approach to Health Behaviors With Interchangeable Elements, Illustrated Through Multiple Sun Protection Behaviors

Yuelin Li a,b, Elizabeth Schofield a & Jennifer L. Hay a

a Department of Psychiatry and Behavioral Sciences
b Department of Epidemiology and Biostatistics, Memorial Sloan Kettering Cancer Center

Abstract Analytics for behavioral habit typically model one behavior at a time, despite the fact that habit often involves multiple cooccurring behaviors, such as food choices and physical activities, where interrelated behaviors are often equally recommended. We propose a novel Mixed-Effects Dynamic hAbit model (MEDA) to simultaneously model multiple related, habitual behaviors. As an illustrative example, MEDA was applied to real-time assessments of sun protection (sunscreen, shade, hat, and protective clothing) reported twice daily by first-degree relatives of melanoma patients who are themselves at an elevated risk of skin cancer. MEDA aims to explicate habits in sun protection under varying environmental cues (e.g., sunny and hot weather). We found consistent between-group differences (e.g., men responding to weather cues more consistently than women) and interactions between cooccurring behaviors (e.g., being in shade discourages sunscreen wearing, more so in men than women). Moreover, MEDA transcends conventional methods to address longstanding challenges—how cue to action and volitional choices differ by groups or even by individual persons. Such nuances in interrelated habitual behaviors are relevant in numerous other applications, such as recommended dietary or physical activity behaviors. These methods best inform personalized behavioral interventions targeting individual preferences for precision behavioral intervention.

Keywords habit; sun protection; Bayesian hierarchical models; variance heterogeneity; conditional probability.

Introduction

The growing literature on the theory of habit characterizes habit as originating first from deliberative, goal-oriented, slower and more cognitively involved processes (Gardner et al., 2014; Neal et al., 2012, 2006; Wood & Neal, 2007; Wood & Runger, 2016). Through repetition and practice, the cognitive association is strengthened and increasingly relegated into fast and intuitive mental short-cuts that appear to involve minimal conscious awareness (Gardner, 2015; Gardner et al., 2014; Hagger, 2019; Marien et al., 2018; Phillips, 2019; Wood & Neal, 2009). Once acquired, habits are often activated spontaneously, as if they are left on autopilot to an outside observer and instigated by a cue to action. Many activities fit this characterization, for instance, getting into a car activates putting on the seatbelt. Going to the movies may activate the purchase of popcorns (Wood & Runger, 2016). As such, habit formation may be viewed as a gradual transition from an initially effortful action to an automatic response to cues. Repetitions strengthen the cue-target association, so that conscious remembering, planning and enactment become increasingly unnecessary. These key characteristics motivate Gardner’s “habit-formation model” (Gardner et al., 2014), which emphasizes repetition in stabilizing habits in behavioral change interventions (Danner et al., 2008; Hagger, 2019; Lally et al., 2010; Marien et al., 2018; Phillips, 2019; Wood & Neal, 2009). Once acquired, habits are often activated spontaneously, as if they are left on autopilot to an outside observer and instigated by a cue to action. Many activities fit this characterization, for instance, getting into a car activates putting on the seatbelt. Going to the movies may activate the purchase of popcorns (Wood & Runger, 2016). As such, habit formation may be viewed as a gradual transition from an initially effortful action to an automatic response to cues. Repetitions strengthen the cue-target association, so that conscious remembering, planning and enactment become increasingly unnecessary. These key characteristics motivate Gardner’s “habit-formation model” (Gardner et al., 2014), which emphasizes repetition in stabilizing habits in behavioral change interventions (Danner et al., 2008; Hagger, 2019; Lally et al., 2010; Marien et al., 2018; Phillips, 2019; Wood & Neal, 2009). Once acquired, habits are often activated spontaneously, as if they are left on autopilot to an outside observer and instigated by a cue to action. Many activities fit this characterization, for instance, getting into a car activates putting on the seatbelt. Going to the movies may activate the purchase of popcorns (Wood & Runger, 2016). As such, habit formation may be viewed as a gradual transition from an initially effortful action to an automatic response to cues. Repetitions strengthen the cue-target association, so that conscious remembering, planning and enactment become increasingly unnecessary. These key characteristics motivate Gardner’s “habit-formation model” (Gardner et al., 2014), which emphasizes repetition in stabilizing habits in behavioral change interventions (Danner et al., 2008; Hagger, 2019; Lally et al.,
Vol. 19 (2023) and summary scores (Schofield et al., 2019) – relying on analyzing individual sun protection behaviors (Hay et al., 2017) and summary scores (Schofield et al., 2019). HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components. HLM and GEE tools are widely available and easy to use, and we have applied HLM in analyzing cooccurring components.

Motivated by these new Bayesian approaches, we propose a Mixed-Effects Dynamic hAbit model (MEDA) to examine individual differences in cue responses across multiple behaviors so that it can 1) simultaneously model interchangeable/cooccurring behaviors; 2) incorporate predictors of individual idiosyncrasies in cue-action associations (e.g., identify shade-seekers from hat wearers); and 3) capture complexities between behaviors (e.g., sunscreen use if already in shade and wearing a hat). Sun protection behaviors are used as an illustrative example.

The paper is organized as follows. We first describe the background rationale to capture sun protection behaviors of first-degree relatives of melanoma patients, who are at an increased risk of developing melanoma themselves (Ford et al., 1995). Then we first apply the model to a single behavior of sunscreen use. Next, the model is extended to simultaneously model inter-dependencies across multiple sun protection behaviors. A portion of the raw data is provided in the Appendix so that readers can immediately...
Table 1: Observed frequency of sunscreen use between males and females.

<table>
<thead>
<tr>
<th></th>
<th>Male</th>
<th></th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Sunny &amp; hot</td>
<td>Not</td>
<td>Sunny &amp; hot</td>
</tr>
<tr>
<td>sunscreen Yes</td>
<td>169</td>
<td>103</td>
<td>234</td>
</tr>
<tr>
<td>sunscreen No</td>
<td>170</td>
<td>31</td>
<td>364</td>
</tr>
</tbody>
</table>

*Note. N = 1,304 entries altogether.*

apply the example analysis to verify the results.

**Methods**

**Sample on Sun Protection Against Familial Risk of Melanoma**

Study design and human subject research ethics were reviewed and approved by the Institutional Review Board of our institution. We enrolled 59 first-degree relatives of melanoma patients. Upon enrollment, participants chose a 14-day interval in the summer months when outdoor sun exposure was anticipated every morning and afternoon (1,312 total entries), a pragmatic consideration aimed to reduce concerns over environmental factors that mitigated the need for sun protection (e.g., participant being indoors at work or at home). An interactive voice response system (IVRS) called the participants’ cell phones and requested keypad responses (y/n) on real-time sun protection behaviors (sunscreen, shade, hats, protective clothing) and decision factors (weather, type of activity, convenience, social support), made twice daily (morning and afternoon). There were altogether 28 assessments per participant. There was minimal missing data (Holland et al., 2020). Additional details on recruitment and assessment procedures can be found elsewhere (Hay et al., 2017). Table 1 tallies up responses on sunscreen use, stratified by the participants’ sex and participant-reported sunny and hot weather. The combined ‘sunny and hot day’ in the assessments, rather than ‘sunny’ and ‘hot’ days separately, was based on our previous qualitative interviews with first-degree relatives of melanoma patients (Shuk et al., 2012). Interviewees systematically used ‘sunny and hot’ together to characterize weather cues that prompted sun protection behaviors.

**Single Outcome Model on Sunscreen Use**

Consider a study of n participants on sunscreen use y_{i,t} from the ith person at assessment time t, where the bracketed index i [t] denotes that the assessments are nested within study participants. We wish to model the probability of sunscreen use as a function of sunny and hot weather in a two-level model (Raudenbush & Bryk, 2002):

**Level 1:**

\[ \Pr(y_{i,t} = "yes") = \logit^{-1}(\beta_{0i} + \beta_{1i} \cdot \text{sunny.hot}_{i,t}) \]

**Level 2 (person level):**

\[ \beta_{0i} = \gamma_{00} + \gamma_{01} \text{male}_i + u_{0i}, \]

\[ \beta_{1i} = \gamma_{10} + \gamma_{11} \text{male}_i + u_{1i}, \]

\[ \begin{bmatrix} u_{0i} \\ u_{1i} \end{bmatrix} \sim \mathcal{N}(0, \Sigma) \]

\[ \Sigma = \Sigma_f, \quad \text{when sex = female}, \]

\[ \Sigma = \Sigma_m, \quad \text{when sex = male} \]

\[ \Sigma_f = \begin{bmatrix} \sigma^2_{0f} & \rho_f \sigma_{0f} \sigma_{1f} \\ \rho_f \sigma_{0f} \sigma_{1f} & \sigma^2_{1f} \end{bmatrix}, \]

\[ \Sigma_m = \begin{bmatrix} \sigma^2_{0m} & \rho_m \sigma_{0m} \sigma_{1m} \\ \rho_m \sigma_{0m} \sigma_{1m} & \sigma^2_{1m} \end{bmatrix}. \]

In level 1, each person’s probability of sunscreen use over time is a function of weather, on not sunny and hot days (\(\beta_{0i}\), intercept for the ith person) and sunny and hot days (slope \(\beta_{1i}\)). The intercepts and slopes are further analyzed in level 2 to yield the following fixed effects:

\(\gamma_{00}\): overall log odds of sunscreen use for women on not sunny and hot days.

\(\gamma_{10}\): for women, the change in sunscreen use when sunny and hot.

\(\gamma_{01}\): difference in sunscreen use between men and women when not sunny and hot.

\(\gamma_{11}\): The difference in sunny and hot response between men and women.

**Random effects as indicators of habit and response to cues.** The random intercepts \(u_{0i}\) and random slopes \(u_{1i}\) represent, respectively, each person’s unique sunscreen use, above and beyond the group (i.e., sex) averages in the fixed effects. The random effects are assumed to follow a bivariate normal distribution with mean 0 and covariance matrix \(\Sigma\). We include separate random effects covariances for women (\(\Sigma_f\)) and men (\(\Sigma_m\)). The standard deviations \(\sigma_{0f}\) and \(\sigma_{0m}\) capture women’s and men’s consistencies in habit in cool weather, where a lower standard deviation indicates a higher consistency. Similarly, \(\sigma_{1f}\) and \(\sigma_{1m}\) capture the consistencies in their responses to the cue of sunny and hot weather. The correlations \(\rho_{0f}\) and \(\rho_{m}\) capture the extent to which participants respond to weather if he or she is already using sunscreen regularly in cool weather.

A central contribution of the MEDA model is in the heterogeneous random effects \(\Sigma_f\) and \(\Sigma_m\), rather than the default homogeneous assumption in standard HLM. We use...
the 2-level notation of Raudenbush and Bryk (2002) because it is probably familiar to social and behavioral scientists. The notation can be easily extended to a full Bayesian model expression (e.g., Gelman et al., 2013, Chapter 14; and McElreath, 2020, section 4.4.1). Details on the priors used are outlined in the Supplementary Materials.

Hypotheses on model parameters. The varying consistencies provide a way to address research questions not easily addressable by conventional tools. As an illustrative example, suppose a researcher hypothesizes that men are more consistent than women in responding to hot and sunny weather, which may be expressed thusly:

\[ H : \sigma_{1f} < \sigma_{1m} \]
\[ H^c : \sigma_{1f} \geq \sigma_{1m} \]

where after collecting and rearranging the terms, the final equation contains fixed effects for the male sex (\(y_{01}\)), sunny and hot weather (\(y_{10}\), and an interaction between the two (\(y_{11}\)). The random effects, \([u_{0i} + u_{1i} \cdot \text{sunny.hot}_{i[t]}]\) can be easily incorporated into model syntax (see Supplementary, sections 1 – 2).

LKJ prior on the random effects. The Cholesky factorization is applied to the covariance matrices \(\Sigma_f\) and \(\Sigma_m\) into diagonal matrix of standard deviation and a correlation \(R\):

\[ \Sigma = \begin{bmatrix} \sigma_0 & 0 \\ 0 & \sigma_1 \end{bmatrix}, \quad R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix} \]

and the correlation \(R\) follows an LKJ prior (Lewandowski et al., 2009),

\[ R \sim \text{LKJCorr}(2) \]

and the LKJ prior offers more flexibility than the inverse-Wishart prior (Hipson, 2020; MRC Biostatistics Unit, 2020; Thompson, 2014).

Bayesian workflow and prior specification. We follow the recommended Bayesian analysis workflow (Depaoli & van de Schoot, 2017), which begins with prior specifications using the recommended “weakly informative” priors (Gelman, 2007), e.g., \(\mathcal{N}(0, 2.5)\) for the fixed coefficients and Cauchy(0, 10) for the intercept (Gelman et al., 2008). Sensitivity to priors is evaluated by replacing the weakly informative priors with vague but proper prior (e.g., \(\mathcal{N}(0, 10^5)\)). Model assumptions are evaluated by the Bayes factor (BF Kass & Raftery, 1995) including the main assumption of separate random effects for men and women. Bayes factor is calculated using the bridgesampling package in R (Gronau et al., 2020) which implements the bridge sampling algorithm to compare non-nested models (Good, 1988; Jackman, 2009). All models are fitted with the Stan programming language (Stan Development Team, 2021) via the rstan package in R (R Core Team, 2021, syntax code in Supplement). Markov chain simulations generally involves four chains of 25,000 iterations each, with a warmup of 5,000 iterations and a thinning interval of 10, which yield a total of 12,000 samples from the posterior draws. Convergence of the simulations is evaluated by the \(\hat{R} \leq 1.01\) diagnostic metric (Gelman & Rubin, 1992; Stan Development Team, 2020). Parameter estimates and their 95% Highest Density Intervals (HDI) are sought.

Modeling Multiple Sun Protection Behaviors Simultaneously

Equation (1) can be applied to each of the four sun protection behaviors:
where the four equations can be combined into one with four behaviors indexed by $k$:

$$
\text{Pr}(y_{i[t]} = \text{"yes"}) = \logit^{-1} \left( \gamma_{00} + \gamma_{01} \text{male}_i + \gamma_{10} \text{sunny.hot}_{i[t]} + \gamma_{11} \text{male}_i \cdot \text{sunny.hot}_{i[t]} + [u_{0i} + u_{1i} \cdot \text{sunny.hot}_{i[t]}] \right)
$$

where $k = 1$ (sunscreen), 2 (shade), 3 (hat), and 4 (long sleeve). The random effects $u_{ik}$ represent four pairs of intercepts and slopes with a covariance $\Psi$, one slice for women and the other for men (but not separately denoted to simplify the notation). Details on how to implement this model in the Stan language is summarized in the Supplement, section 5.

**Hypotheses on the Dynamics Between Correlated Sun Protection Behaviors.**

Recall the hypotheses on cooccurring behaviors, that we want to capture complexity such that a person may skip sunscreen if he or she is already in the shade and wearing a hat, which can be expressed mathematically as a conditional probability $p(\text{sunscreen}|\text{hat, shade})$. We can derive this conditional probability for men and women separately. Also, if $p(\text{sunscreen}|\text{hat, shade})$ is substantially smaller than $p(\text{sunscreen})$, this is an indication that sunscreen use is suppressed by shade and hat usage. The Supplement contains illustrative examples on how to leverage this feature to evaluate behavioral dynamics.

**Results**

**Participant Characteristics**

Participant characteristics have been described previously (Hay et al., 2017). Briefly, the participants were non-Hispanic Whites (98%), 64% female, 74% with a college degree or above, and 62% married/partnered. All had a first-degree family history of melanoma.

**Single-Outcome Model Results**

Table 2 summarizes the fitted model in sunscreen use. Men are less likely to use sunscreen when the weather is not sunny and hot ($\gamma_{10}$: Odds Ratio = 0.18, 95% HDI: 0.02, 1.28). Women’s sunscreen use is reliably greater in sunny and hot weather ($\gamma_{11}$: OR = 6.14, 95% HDI: 3.38, 11.24). Men are comparatively more responsive to sunny and hot weather ($\gamma_{10}$: OR = 4.73, 95% HDI: 0.79, 34.48; 90% HDI: 1.04, 22.43), although the HDI only excludes the null with 90% posterior probability.

**Within-person variations in sunscreen use.**

Table 2 shows that women’s response to sunny and hot weather is more consistent than men’s, shown in the smaller standard deviation ($\sigma_{1f} = 0.76$ vs. $\sigma_{1m} = 2.52$). The
Parameter estimates for the full model (1) in the single outcome of sunscreen use.

<table>
<thead>
<tr>
<th>Fixed parameters</th>
<th>Variable</th>
<th>Coefficient</th>
<th>Exp(Coefficient)</th>
<th>95% HDI *</th>
</tr>
</thead>
<tbody>
<tr>
<td>γ₀₀</td>
<td>(Intercept)</td>
<td>-1.27</td>
<td>0.28</td>
<td>0.11, 0.66</td>
</tr>
<tr>
<td>γ₀₁</td>
<td>male</td>
<td>-1.74</td>
<td>0.18</td>
<td>0.02, 1.28</td>
</tr>
<tr>
<td>γ₁₀</td>
<td>sunny.hot</td>
<td>1.82</td>
<td>6.14</td>
<td>3.48, 11.24</td>
</tr>
<tr>
<td>γ₁₁</td>
<td>male · sunny.hot</td>
<td>1.55</td>
<td>4.73</td>
<td>0.79, 34.48</td>
</tr>
</tbody>
</table>

Random effects

<table>
<thead>
<tr>
<th></th>
<th>σ₁⁻ᶠ</th>
<th>σ₂⁻ᶠ</th>
<th>ρ⁻ᶠ</th>
<th>σ₁⁻ᵐ</th>
<th>σ₂⁻ᵐ</th>
<th>ρ⁻ᵐ</th>
</tr>
</thead>
<tbody>
<tr>
<td>[γ₁⁻ᶠ]</td>
<td>σ₁⁻ᶠ = 2.19</td>
<td>[1.44]</td>
<td>[3.01]</td>
<td>[σ₁⁻ᵐ]</td>
<td>σ₁⁻ᵐ = 3.52</td>
<td>[1.73]</td>
</tr>
<tr>
<td>[σ₁⁻ᵐ]</td>
<td>[σ₁⁻ᶠ]</td>
<td>0.19</td>
<td>-0.81</td>
<td>0.12</td>
<td>-0.52</td>
<td>0.69</td>
</tr>
<tr>
<td>[σ₂⁻ᵐ]</td>
<td>[σ₂⁻ᶠ]</td>
<td>[σ₂⁻ᵐ]</td>
<td>0.53</td>
<td>4.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>[ρ⁻ᵐ]</td>
<td>[ρ⁻ᶠ]</td>
<td>[ρ⁻ᵐ]</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Note. *: HDI stands for Highest Density Interval of the posterior distribution.

posterior probability in \( \Pr(σ₁⁻ᶠ < σ₁⁻ᵐ|γ_{₁|n}) \) is 0.95, which is calculated from the MCMC draws of the posterior distributions (see Supplement, section 6). The Bayes factor is therefore the posterior odds of \( 0.95/(1–0.95) = 9.5 \), a 'strong' evidence (Kass & Raftery, 1995) in favor of the hypothesis. Similarly, in cool weather, women are also more consistent than men \( (σ₁⁻ᶠ = 2.19 \text{ vs. } σ₀⁻ᵐ = 3.52) \), which yields a posterior probability of 90% and a 'strong' evidence.

Table 2 also shows that \( ρ⁻ᶠ = -0.19 \), the correlation between women's random intercepts and slopes is -0.19. Women who use sunscreen infrequently on cooler days are slightly more likely to use it when it is sunny and hot. In men, the correlation is a positive 0.12, a weak correlation between men's sunscreen use in cooler days and sunny and hot days. However, none of the posterior highest density regions exclude the null, suggesting a high degree of uncertainty in these correlations.

**Individual variabilities in the habit of sunscreen use.** Figure 1 plots the idiosyncratic habits of individual persons to further elucidate the findings in Table 2. Model-estimated probability of sunscreen use is plotted against the observed average sunscreen use when not sunny and hot. Also plotted are the changes in sunny and hot weather to display each person's unique response to weather cue. There appears to be a cluster of participants (lower left corner) who rarely use sunscreen in cool weather. However, some of them respond to sunny and hot weather with an upward of 60% to 70% sunscreen use. The spread in Figure 1 also shows considerable variabilities in the habit of sunscreen use. It appears that almost all women respond to sunny and hot weather, whereas some men hardly move. This observation corroborates the last hypothesis in Table 2, where women are more consistent than men in responding to sunny and hot weather.

Figure 2 plots the random slopes \( (u_{₁|l}) \) against random intercepts \( (u_{₀|l}) \) derived from posterior means. Each symbol represents one person. The error bars suggest that there is uncertainty in each person's slope and intercept.

**Multiple Sun Protection Behaviors Modeled Simultaneously**

Table 3 at the end of the article shows the parameter estimates for model (2) with all four behaviors modeled simultaneously. An estimated 24% of women use sunscreen in cooler weather (estimated odds, \( γ₀|\text{sunscreen} = 0.31, 95\% \text{ HDI: 0.15, 0.61} \)). On cool days, women are discernably less likely to wear a hat than use sunscreen \( (OR = 0.31, 95\% \text{ HDI: 0.11, 0.77}) \) but more likely to wear long sleeve clothing \( (OR = 3.85, 1.32, 11.85) \). By comparison, men are less likely to use sunscreen \( (γ₀|\text{sunscreen} = 1.30, OR=0.28, 0.08, 1.06) \) but much more likely to wear a hat \( (γ₀|\text{hat}=2.48, OR=11.65, 2.05, 69.43) \). Sunny and hot days prompt both sexes to use sunscreen, for women \( γ₀|\text{sunscreen} = 1.73, OR=5.60, 3.25, 9.90) \) and men \( γ₁|\text{sunscreen} = 1.03, OR=2.77 \) as compared to women, 0.88, 8.58). The OR of 2.77 indicate that men are more responsive to sunny and hot weather than women at a posterior confidence of 80%. Long sleeve clothing is unpopular on sunny days for both women \( (γ₁|\text{sleeve} = -2.58, OR=0.07, 0.03, 0.19) \) and men \( γ₁|\text{sleeve} = -0.10, OR=0.93, 0.18, 4.70) \). Hat wearing and shade seeking are equally likely than sunscreen use on sunny days (neither \( γ₁|\text{hat} \) nor \( γ₁|\text{shade} \) exclude null).

Also shown in Table 3 are the random effects correlation matrices \( Ψ_f \) and \( Ψ_m \). It shows greater variabilities in men than women, e.g., in sunscreen use (0.85 probability that \( σ₀⁻ᵐ = 2.32 \geq σ₀⁻ᶠ = 1.79 \). Recall that Bayes factor in \( H : σ₀⁻ᶠ < σ₀⁻ᵐ \) is fully determined by the posterior odds of \( 0.85/(1–0.85) = 5.67 \), a ‘substantial’ evidence in favor of the hypothesis. Men also show a greater variability than women in increased sunscreen use in response to sunny and hot weather (0.84 posterior probability that \( σ₀⁻ᵐ = 1.38 \geq σ₀⁻ᶠ = 0.87 \), also a ‘substantial’ posterior odds.
Figure 1. Model-estimated sunscreen for individual persons. Each point represents one person. Small amounts of random displacements are added to the points to better separate them. Error bars represent the 50% posterior predictive intervals. To facilitate a clearer visual comparison, the filled symbols represent the model-estimated sunscreen use when the weather is sunny and hot, by matching them to the same person’s observed use when it is not. The vertical increase is visibly greater in men than in women. There are considerable individual variabilities, e.g., participant 7006 (marked with arrows) was strongly influenced by the weather. Several men’s sunscreen use was below 20% even in sunny and hot weather.

Average differences between the sexes. Figure 3 displays the model-estimated probability of sun protection behaviors. Women frequently use long sleeve clothing on not sunny and hot days. On sunny and hot days, women rely less on long sleeve clothing and hat and turn to sunscreen and shade. Men most often wear a hat on not sunny and hot days.

Conditional probabilities between cooccurring behaviors. Table 4 at the end of the article summarizes the conditional probabilities between co-occurring behaviors in response to sunny and hot weather, stratified by sex. The pair between $p($sunscreen$|hat)$ and its reverse probability $p($hat$|sunscreen)$ allows a comparison between the order of adding sun protections to one that is already in use (e.g., whether or not adding hat to sunscreen is easier than the other way around).

For men in cooler weather, adding hat to sunscreen is slightly easier (0.39) than the reverse order (0.27). For women, however, the opposite is found. Adding sunscreen to hat is easier (0.48) than the reverse (0.27). Sunny and hot weather does not affect this preference, as women continue to prefer adding sunscreen to hat (0.70 vs. 0.37), while men prefer the reverse (0.67 vs. 0.59). The non-overlapping HDIs indicate statistically reliable differences. No strong preference is found in the other two pairs of conditional probabilities. Finally, women engage in adding one other sun protection in addition to sunscreen at a probability of 0.41 on cooler days (0.32 in men); this difference disappears on sunny and hot days (0.54 vs. 0.47 for women and men, respectively).

More complex combinations such as $p($sunscreen$|hat,shade)$ show 0.39 and 0.71 for women and 0.0 and 0.52 for men on cool and sunny days, respectively (see Supplement, section 7, for detailed calculation). Men on cool days skip sunscreen entirely when they are already in shade and wearing a hat whereas 61% of women do the same; still 48% of men skip sunscreen on sunny and hot days given hat and shade.

Within-person differences in response to weather. Figure 4 plots the estimated probability of four different sun protection behaviors, stratified by sunny and hot weather. Each circle represents one person’s four behaviors. Overall, considerable individual differences are visible in their choices of sun protection, even on not sunny and hot days. Some individuals use sunscreen often but do not seek shade. On sunny and hot days, there is a visible move-
Figure 2  Random slopes ($u_{1i}$) plotted against random intercepts ($u_{0i}$). Each symbol represents one person. The random slopes represent each person’s sunscreen use when the weather is sunny and hot, centered at women’s average. The random intercepts represent sunscreen use when the weather is not sunny and hot, also centered at women’s average. Error bars represent the 50% posterior highest density regions of the random slopes (not 90% to minimize visual clutter). Men’s sunscreen use is visibly affected by weather (larger concentration of blue dots with a slope above zero).

Discussion

The proposed Mixed-Effects Dynamic hAbit model (MEDA) explains individual variations in their sun protection behaviors. MEDAs’ central contributions is an analytic framework more flexible than the conventional EMA analytics on one outcome at a time (Schwartz & Stone, 1998). By analyzing multiple cooccurring behaviors simultaneously, and by allowing subgroups to have unique and separate within-group covariances, the proposed MEDA model addresses research questions unanswerable by conventional approaches. For instance, individuals who already have a habit to seek shade are less likely to use sunscreen, and these associations are explicitly defined as unique covariances for subgroups. Furthermore, adding sunscreen to hat wearing is easier for women than for men, but the opposite is found in adding a hat to sunscreen, possibly attributable to the base rate habits of what men and women typically use and how they typically respond to the weather cue. These variations are often treated as noise to be discarded in conventional methods, but MEDA capitalizes on them to address complex research questions not easily addressable by traditional methods. Note that sex differences are illustrative examples, and the analytic framework can be easily extended to other stratification factors (e.g., age, educational attainment). The proposed new statistical approach addresses longstanding challenges in nearly all observational studies of health behaviors, and is applicable in numerous other habits, such as nutrition, physical activity, and medication adherence.

Individual variation plays an important role in both the theory of habit and in intervention development. Researchers often hold a theoretical postulation that habit, once acquired and maintained, require nearly no cognitive effort to enact (Gardner, 2015; Hagger, 2019; Phillips, 2019; Wood & Neal, 2007, 2009; Wood & Runger, 2016). However, our findings are inconsistent with this theoretical postulation. For instance, the majority of men on cool days skip sunscreen entirely when they are already in shade and wearing a hat, whereas only 61% of women do the same. Additionally, the conditional probabilities suggest that habit maintenance remains stochastic depending on the circumstances, and cognitive efforts may still be needed to evaluate the environmental cues and other circumstances of sun protection. Thus, habit in the specific context of sun protection may not be as consistent and un-
changing as postulated. Automatic enactment of habit may not be universally true in all health behaviors. That is why we need a versatile tool like MEDA to allow researchers in diverse areas of health behaviors to capture the inherently stochastic actions and their cognitive antecedents, while simultaneously allowing subgroups to have idiosyncrasies. It is precisely in these within-person variabilities that targeted interventions may be most effective.

These nuanced findings inform why a one-size-fits-all best advice in sun protection may no longer be enough. The International Agency for Research on Cancer (2001) recommends the use of sunscreen plus at least one other method (shade, hat, and protective clothing), which suggests that one should adopt as many of these protections as possible into one's lifestyle and make them daily habits regardless of weather and season. However, our analysis on individual variabilities and the conditional probabilities show that such a general advice is unlikely to work for all. Targeted interventions taking into consideration existing personal preferences may be essential, as shown in our findings that adding extra sun protection depends on what the person is already using, and men and women show consistently different preferences. In short, heterogeneity in response to cues found in a statistical model may be leveraged for a personalized intervention plan.

There are limitations in the proposed model. Specific to sun protection behaviors, we have not exhausted all options (e.g., sunglasses, transition lenses, visors and other protective clothing options), predictors (e.g., convenience, social support), and subgroups (e.g., racial groups, geographical locations, first-degree relatives of melanoma patients vs. the general public). A study participant may be more likely to put on sunscreen if it is readily available and easy to apply (e.g., spray sunscreens). Shade seeking may be more likely if it is accessible (e.g., trees in a park, umbrellas on a beach) and supported by others (e.g., adults setting up an umbrella for children). Variabilities between other stratification factors such as racial groups may also be important. Note, however, that MEDA can be easily extended to include more concurrent behaviors, covariates, and subgroups. For instance, convenience, social support, and weather cues may play synergistic roles in sun protection behaviors. If sunny weather is antici-
Figure 4  Estimated probability of sun protection behaviors, stratified by weather. Each circle represents one person’s four behaviors. Women are plotted in pink and men in blue. The probabilities of applying sunscreen and shade seeking are plotted on the x and y-axes, respectively. Wearing a hat and long sleeve clothing are represented by circle sizes and color saturation, respectively, as indicated in the legend. The left panel shows a pattern that, on not sunny & hot days, some individuals use sunscreen often but do not seek shade and vice versa. Others mostly seek shade and wear a hat or long sleeve clothing. Few do all four, and they are more likely to be women than men. On the right panel, grey arrows are added to indicate changes in behaviors when the weather changes. The most visible pattern shows a general movement from lower left to upper right, with larger circles and saturated color, indicating an overall increase in all sun protection behaviors in most individuals. Some individuals appear to move up (more shade seeking, but other behaviors relatively unchanged). Another subset of individuals appear to move to the right (more sunscreen mostly).

pated but shade may not be conveniently available, then social support may become key to adequate sun protection (e.g., adults bringing sunscreens, hats, sunglasses, and beach umbrellas for children). Future research may include predictors such as interaction effects between decision factors to address such research questions. One possible concern regards the size of stratification groups; we had fewer male than female participants and variability tends to be greater in smaller subgroups (thus possible statistical artefact in the hypotheses comparing variances). There is also a concern regarding resources and participant burden; the model in its current form is tailored to intensive longitudinal measures, and thus is more resource intensive than conventional designs that yield only a handful of longitudinal assessments. However, our twice-daily assessments demonstrate that intensive longitudinal measurement can be done economically over participants’ cellular phones. Another concern is that intensive longitudinal assessments may become cues for habit. However, a careful examination shows that this reminder effect is unlikely (Schofield et al., 2019).

Future research may also extend beyond sun protection to other areas of habit research such as physical activity, substance use, and nutrition, where multiple interchangeable behaviors may be involved but typically analyzed individually in current applications. Moreover, the statistical properties of the proposed method can also be examined more systematically, such as sensitivity to priors, particularly on sparse sun protection options. Another more technical concern is that programming MEDA is complex and not immediately available to researchers inexperienced with these tools. However, the syntax recipes and the raw data in the Appendix should help readers practice model fitting. Previous applied studies often examine elements of a habit individually, despite the knowledge that they may cooccur and one specific element may affect the enactment of other interchangeable elements. Our MEDA approach offers a flexible and scalable framework to understand more fully the complex interactions between multiple cooccurring behaviors in a habit; and it maps out between-group averages and heterogeneity attributable to individual persons such as the more volitional aspects of habit. Our previous, one-behavior-at-a-time approach (Hay et al., 2017) did not include a correlation between multiple behaviors. Therefore, we had no access to some of the new findings reported herein: 1) shade seeking...
suppresses sunscreen use in both men and women; 2) conditional probability in hat wearing if someone is already in shade is different between men and women; and 3) subtle differences in the correlations between behaviors in men and women. Indeed, the conventional approach clearly lacks the ability to detect such nuanced but important differences, which may shed light on the development of future behavioral interventions tailored to individual differences. There may not be one-size-fits-all interventions, a personalized intervention may be needed to tailor to an individual person’s existing preferences. Parsing out individual differences from averages is key to a more tailored and personalized intervention that can be broadly applied to numerous other habit behaviors.

Authors’ note

Acknowledgments of grants and funding. This work was in part funded by the National Institute of Health grants R21 CA137532 (JH, Principal Investigator) and NIH Support Grant P30 CA008748 to Memorial Sloan Kettering Cancer Center.

References


Goldstein, S. P., Dochat, C., Schumacher, L. M., Manasse, S. M., Crosby, R. D., Thomas, J. G., Butryn, M. L., & For-


Open practices

The Open Material badge was earned because supplementary material(s) are available on the journal’s web site.

Citation


Copyright © 2023, Li et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) or licensor are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Received: 07/02/2023 ~ Accepted: 14/08/2023

Tables 3 and 4 follows.
Table 3: Parameter estimates on four sun protection behaviors modeled simultaneously.

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Variable</th>
<th>Coefficient</th>
<th>Exp(Coefficient)$^a$</th>
<th>95% HDI $^b$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\gamma_{00}[\text{sunsunreact}]$</td>
<td>(Intercept)</td>
<td>-1.19</td>
<td>0.31</td>
<td>0.15  0.61</td>
</tr>
<tr>
<td>$\gamma_{01}[\text{sunsunreact}]$</td>
<td>Male</td>
<td>-1.30</td>
<td>0.28</td>
<td>0.08  1.06</td>
</tr>
<tr>
<td>$\gamma_{00}[\hat{}]$</td>
<td>Hat</td>
<td>-1.16</td>
<td>0.31</td>
<td>0.11  0.77</td>
</tr>
<tr>
<td>$\gamma_{00}[\text{shade}]$</td>
<td>Shade</td>
<td>0.10</td>
<td>1.11</td>
<td>0.47  2.63</td>
</tr>
<tr>
<td>$\gamma_{00}[\text{sleeve}]$</td>
<td>Long Sleeve</td>
<td>1.35</td>
<td>3.85</td>
<td>1.32  11.85</td>
</tr>
<tr>
<td>$\gamma_{10}[\text{sunsunreact}]$</td>
<td>Sunny · Hot</td>
<td>1.73</td>
<td>5.60</td>
<td>3.25  9.90</td>
</tr>
<tr>
<td>$\gamma_{01}[\hat{}]$</td>
<td>Male · Hat</td>
<td>2.48</td>
<td>11.65</td>
<td>2.05  69.43</td>
</tr>
<tr>
<td>$\gamma_{00}[\text{shade}]$</td>
<td>Male · Shade</td>
<td>0.89</td>
<td>2.34</td>
<td>0.34  14.89</td>
</tr>
<tr>
<td>$\gamma_{01}[\text{sleeve}]$</td>
<td>Male · Sleeve</td>
<td>-0.54</td>
<td>0.58</td>
<td>0.05  4.77</td>
</tr>
<tr>
<td>$\gamma_{11}[\text{sunsunreact}]$</td>
<td>Male · Sunny</td>
<td>1.03</td>
<td>2.77</td>
<td>0.88  5.58</td>
</tr>
<tr>
<td>$\gamma_{10}[\hat{}]$</td>
<td>Hat · Sunny</td>
<td>-0.72</td>
<td>0.49</td>
<td>0.21  1.07</td>
</tr>
<tr>
<td>$\gamma_{10}[\text{shade}]$</td>
<td>Shade · Sunny</td>
<td>-0.27</td>
<td>0.77</td>
<td>0.36  1.65</td>
</tr>
<tr>
<td>$\gamma_{10}[\text{sleeve}]$</td>
<td>Sleeve · Sunny</td>
<td>-2.58</td>
<td>0.07</td>
<td>0.03  0.19</td>
</tr>
<tr>
<td>$\gamma_{11}[\text{sunsunreact}]$</td>
<td>Male · Hat · Sunny</td>
<td>-0.54</td>
<td>0.59</td>
<td>0.15  2.16</td>
</tr>
<tr>
<td>$\gamma_{11}[\text{shade}]$</td>
<td>Male · Shade · Sunny</td>
<td>-1.00</td>
<td>0.38</td>
<td>0.08  1.45</td>
</tr>
<tr>
<td>$\gamma_{11}[\text{sleeve}]$</td>
<td>Male · Sleeve · Sunny</td>
<td>-0.10</td>
<td>0.93</td>
<td>0.18  4.70</td>
</tr>
</tbody>
</table>

**Random Effects**

<table>
<thead>
<tr>
<th>$\Psi_f$</th>
<th>Not Sunny and Hot Weather</th>
<th>Sunny and Hot Weather</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u_{0}[\text{sunsunreact}]$</td>
<td>1.79</td>
<td>1.79</td>
</tr>
<tr>
<td>$u_{0}[\hat{}]$</td>
<td>-0.23</td>
<td>0.33</td>
</tr>
<tr>
<td>$u_{0}[\text{shade}]$</td>
<td>-0.62</td>
<td>0.26</td>
</tr>
<tr>
<td>$u_{0}[\text{sleeve}]$</td>
<td>-0.18</td>
<td>-0.13</td>
</tr>
<tr>
<td>$u_{0}[\text{sunsunreact}]:\text{hot}$</td>
<td>-0.06</td>
<td>0.06</td>
</tr>
<tr>
<td>$u_{0}[\hat{}]:\text{hot}$</td>
<td>-0.23</td>
<td>-0.03</td>
</tr>
<tr>
<td>$u_{0}[\text{shade}]:\text{hot}$</td>
<td>-0.11</td>
<td>0.23</td>
</tr>
<tr>
<td>$u_{0}[\text{sleeve}]:\text{hot}$</td>
<td>-0.11</td>
<td>0.23</td>
</tr>
</tbody>
</table>

*Note.* $^a$: Exponentiated coefficients represent odds ratios. $^b$: HDI stands for Highest Density Interval of the posterior distribution.
Table 4: Conditional probabilities involving the cooccurrences of sunscreen and another single sun protection behavior.

<table>
<thead>
<tr>
<th>Conditional probability</th>
<th>Men observed</th>
<th>Men model</th>
<th>Men 95% HDI</th>
<th>Women observed</th>
<th>Women model</th>
<th>Women 95% HDI</th>
</tr>
</thead>
<tbody>
<tr>
<td>p(sunscreen</td>
<td>hat)</td>
<td>0.27</td>
<td>0.27</td>
<td>0.13</td>
<td>0.41</td>
<td>0.63</td>
</tr>
<tr>
<td>p(hat</td>
<td>sunscreen)</td>
<td>0.42</td>
<td>0.39</td>
<td>0.22</td>
<td>0.57</td>
<td>0.71</td>
</tr>
<tr>
<td>p(shade</td>
<td>sunscreen)</td>
<td>0.03</td>
<td>0.09</td>
<td>0.00</td>
<td>0.20</td>
<td>0.39</td>
</tr>
<tr>
<td>p(sunscreen</td>
<td>shade)</td>
<td>0.03</td>
<td>0.08</td>
<td>0.00</td>
<td>0.17</td>
<td>0.42</td>
</tr>
<tr>
<td>p(sleeve</td>
<td>sunscreen)</td>
<td>0.48</td>
<td>0.47</td>
<td>0.29</td>
<td>0.66</td>
<td>0.34</td>
</tr>
<tr>
<td>p(sunscreen</td>
<td>sleeve)</td>
<td>0.34</td>
<td>0.34</td>
<td>0.22</td>
<td>0.47</td>
<td>0.43</td>
</tr>
<tr>
<td>p(plus one</td>
<td>sunscreen)a</td>
<td>0.31</td>
<td>0.32</td>
<td>0.24</td>
<td>0.40</td>
<td>0.48</td>
</tr>
</tbody>
</table>

Note. a: The “plus one” label refers to that the person is already using sunscreen, plus at least one other method, defined as the average of \( p(\text{hat|sunscreen}) + p(\text{shade|sunscreen}) + p(\text{sleeve|sunscreen}) \).