Illusory Feelings, Elusive Habits: People Overlook Habits in Explanations of Behavior

Asaf Mazar and Wendy Wood
Department of Psychology, University of Southern California

Abstract

Habits underlie much of human behavior. However, people may prefer agentic accounts that overlook habits in favor of inner states, such as mood. We tested this misattribution hypothesis in an online experiment of helping behavior (N = 809 adults) as well as in an ecological momentary assessment (EMA) study of U.S. college students' everyday coffee drinking (N = 112). Both studies revealed a substantial gap between perceived and actual drivers of behavior: Habit strength outperformed or matched inner states in predicting behavior, but participants' explanations of their behavior emphasized inner states. Participants continued to misattribute habits to inner states when incentivized for accuracy and when explaining other people's behavior. We discuss how this misperception could adversely influence self-regulation.

Habits are key to successful functioning in our day-to-day lives. By automating repeated behaviors, habits allow people to consistently get enough sleep, stay fit, eat healthfully, and study (Galla & Duckworth, 2015). Furthermore, engaging in routine, habitual behaviors is associated with a greater sense of security and meaning in life (Avni-Babad, 2011; Heintzelman & King, 2019). Everyday habits also enable multitasking: In an experience-sampling study, 43% of daily behaviors were performed habitually, in that they were repeated frequently in the same location and typically while participants were thinking about something other than what they were doing (Wood et al., 2002).

Habits are mental associations between contexts and responses that develop as people repeat rewarded responses in a given context (Knowlton & Diedrichsen, 2018). Once habits have formed, context cues automatically activate the repeated response in mind (Mazar & Wood, 2018). Habit associations are separate from the behavior they produce: For example, when one enters a car, the habitual response of wearing a seat belt may be mentally activated regardless of whether one overtly acts on it by buckling up.

Given the many opportunities people have in daily life to observe their own repeated actions, one might expect lay theories to accurately account for habits. In support of this, participants in one study read about an office worker who locked in a colleague by turning the office doorknob counterclockwise (Gershman et al., 2016). Participants demonstrated attribution to habit by placing less blame on the office worker if the worker's home doors also opened counterclockwise, implying that the worker acted out of habit. Also relevant, the *actor-observer effect* suggests that people readily attribute their own behavior (but not others' behavior) to environmental influences (Jones & Nisbett, 1971). Because habits are activated by cues in the environment, people may consequently ascribe their own behavior to habits.

Nevertheless, there is reason to believe that people overlook habit when accounting for their own repeated,

Corresponding Author:
Wendy Wood, University of Southern California, Department of Psychology
Email: wendy.wood@usc.edu

habitual behaviors. People overvalue introspective thoughts, feelings, and emotions in self-judgments (Pronin, 2009), and they interpret actions as intentional by default (Rosset, 2008). Illustrating overattribution to inner states, smokers in one study reported that their smoking was triggered by negative affect, even though in-the-moment affect assessments revealed little association between negative affect and subsequent smoking (Shiffman et al., 1997). In a separate diary study, self-described emotional eaters were not more likely to eat in response to negative emotions (Adriaanse et al., 2011). Finally, participants in another study who had stronger habits reported greater certainty in their behavioral intentions, even though these intentions did not predict their future behavior (Ji & Wood, 2007). Taken together, these findings suggest that people may exaggerate the effect of inner states on behavior while discounting the role of context-cued habit.

In sum, the present research tested a potential tendency to overlook the influence of habits on behavior. Such a bias is important to document, given the many downstream influences of lay theories about behavior (e.g., McFerran & Mukhopadhyay, 2013). For example, such a bias may cause people to ineffectively self-regulate by putting too much weight on regulating inner states (such as mood) and too little weight on self-regulation strategies that may better control habits (such as reducing cues that trigger habitual behavior; e.g., Duckworth et al., 2016).

The Present Research

In two studies, we measured the effects of habits and inner states on a behavior and assessed participants' attributions for that behavior. Our first study was an experiment that orthogonally manipulated habit strength and mood to assess their effects on helping. Our second study used ecological momentary assessment (EMA) to track coffee drinking over a typical week. In both studies, we expected miscalibration of the actual and perceived effects of habit to emerge if participants placed less value on habit than on inner states in their attributions, relative to the actual effects of habit and inner states on behavior.

Both studies were approved by the University of Southern California Institutional Review Board. Preregistration plans for both studies, as well as materials, data, and analytic code, can be accessed on OSF at https://osf.io/5xfsm/.

Study 1

Participants first recalled a happy, sad, or neutral event and then completed a simple, supposedly unrelated

Statement of Relevance

Habits are ubiquitous in daily life, but people may overlook the influence of habits, preferring to explain their behavior using inner states. For instance, they may attribute drinking coffee to feeling tired, even though their consumption is actually driven by habit. In two studies, we tested this tendency to overlook habits in favor of inner states by comparing participants' perceptions with the actual predictors of two behaviors: helping other people in a lab experiment and drinking coffee in daily life. Participants attributed these behaviors to mood and fatigue more than to habit, whereas habit had an equal or stronger influence on actual behavior. With these two behaviors, we demonstrated the bias in a lab study manipulating habit and mood as well as with habits in daily life. To the extent that people overlook habit in this way, they will be ill-equipped to effectively self-regulate habitual behavior.

task that trained them in either a strong or a weak habit of pressing one of two computer keys. Immediately after the task, participants pressed one of the keys to indicate whether they were willing to donate a small amount of time to help the researchers. We expected the habit manipulation to lead participants with strong habits to select the response that mapped onto the extensively trained key. Specifically, our hypotheses compared (a) the actual effects of the mood and habit manipulations on participants' decision to help or not help with (b) participants' attributions for their help.

Orthogonally manipulating habits and inner states does not imply that habits and inner states are always uncorrelated. Indeed, in daily life, habits often align with moods, goals, and feelings. However, once formed, habitual responses are triggered directly by context cues, and inner states provide limited input (Wood et al., 2022).

Method

Power analyses. Power analyses for logistic regression were conducted using the *WebPower* package (Version 0.6; Zhang & Yuan, 2018) in R (Version 4.1.0; R Core Team, 2021). Results suggested that 787 participants were required to achieve 80% power for detecting a difference of 10% in helping behavior between conditions (namely, 45% help in one condition vs. 55% in another). Because this study used a novel habit-formation task, this expected

difference was chosen because it represents a plausible moderate effect size.

Participants. We recruited 808 online participants via Prolific (388 male, 391 female, 15 other, 14 declined to answer; age: $M = 35.17$ years, $SD = 13.32$, range = 18–78). An additional 116 participants were excluded because they did not pass the comprehension check, and an additional 91 participants were excluded because of extreme scores (± 2 *SD* from their mood condition's mean; 76 participants) on the Positive And Negative Affect Schedule (PANAS; Thompson, 2007) or because of slow reaction times (median reaction time > 500 ms; 15 participants). Including all participants in the analyses did not notably alter the results (see Table S2 in Supplemental Material available online).

Procedure. After providing informed consent, participants completed an autobiographical emotional-memory task (Mills & D'Mello, 2014) in which they recalled a happy, sad, or neutral memory in response to the following prompt (the description of each emotion condition varied depending on the memory being induced):

> Recall [an event in your life that made you happy/ an event in your life that made you sad/the last time that you brushed your teeth]. Take some time to really experience the event and the feelings associated with it. When you are ready, describe the event below in your own words. You may use between 5-40 words.

Participants had a minimum of 30 s to write about the scenario in a text box, after which they could proceed with the experiment. As a manipulation check, they then completed items from the PANAS (Thompson, 2007).

For the habit-formation task, participants completed 40 trials in which either the letter "m" or the letter "z" appeared on the screen, and they responded by pressing the corresponding key as quickly and accurately as possible. In the *strong-habit* condition, participants practiced either the left ("z") or the right ("m") response on 36 out of 40 trials (90%) and the alternate response on the remaining four trials (10%). The specific response ("z" or "m") was counterbalanced across participants. In the *weak-habit* condition, participants practiced each response equally (20 trials on each side). This task follows habit-formation procedures in prior research (e.g., Hardwick et al., 2019).

Immediately following the task, a screen displayed the helping request: "Are you willing to complete 40 additional trials (~5 minutes) as a favor to us (without additional compensation)?" "Yes" and "No" responses were mapped onto the same keys used in the habit-formation task ("z" and "m"). The behavioral measure was whether participants agreed or not. For participants in the strong-habit condition, the "No" response was always mapped onto the more heavily practiced key. For participants in the weak-habit condition, "Yes" and "No" responses were randomly assigned to each key. Note that the "No" response was overtrained in the strong-habit condition because the helping request (donating 5 min of participants' time) was taxing and might have spurred participants to exert deliberative control in order to decline it. Participants then answered a comprehension check to test whether they understood the help request, along with additional measures (see below).

Measures.

Positive and negative mood. Participants responded to the prompt "indicate the extent to which you feel the mood below RIGHT NOW." On 9-point scales (1 = *very slightly/not at all* to 9 = *extremely*), participants rated positive emotions ("inspired," "determined," "attentive," "proud," "alert," "active") and negative ones ("upset," "hostile," "ashamed," "nervous," "afraid," "guilty") taken from the PANAS (Thompson, 2007). Positive and negative item ratings were averaged to create a positive-affect score (α = .82) and a negative-affect score (α = .80).

Self-attribution. On percentage scales (0% = *not at all important* to 50% or more = *extremely important*), participants rated the extent to which their decision to help or not was due to habit, "I responded automatically, without thinking," and mood, "My mood at the time (I felt good/bad)." The sum of both answers could range from 0% to 100%. Presentation order of mood and habit was counterbalanced in both the attribution and incentivized measures.

Incentivized other-attribution. To minimize judgment biases, we incentivized participants to provide accurate explanations for other people's behavior. The incentive should minimize the effects of conversational norms regarding plausible or socially acceptable explanations for a behavior. Attributions for other people's behavior are additionally informative because they should be relatively unaffected by self-serving biases that could influence self-attributions. Thus, on percentage scales (0% = *not at all important* to 50% or more = *extremely important*), participants indicated, "How important do you think that the following factors are in determining whether OTHER participants agree or decline to complete additional trials?" Participants then rated habit and

mood, as in the self-attribution measure (above). Accurate ratings (within 5% of the study results) earned a chance to win a $10 bonus.

Habit strength. On 7-point scales (1 = *strongly disagree* to 7 = *strongly agree*), participants rated the extent to which "hitting a key (z or m) in the task is something that I . . ." (a) "did without thinking," (b) "did automatically," (c) "did without having to consciously remember," and (d) "started doing before I realized." These items were taken from the Self-Reported Behavior Automaticity Index (Gardner et al., 2012), which consists of a subset of items from the Self-Report Habit Index (SRHI; Verplanken & Orbell, 2003). Ratings were averaged to create a *perceived-automaticity* score (α = .85).

To our knowledge, this is the first use of this measure with a simple finger-movement task, and experienced automaticity did not differ between the weak-habit (*M* = 4.09) and strong-habit (*M* = 4.09) conditions, 95% CI = [−0.25, 0.24], *t*(479.57) = −0.06, *p* = .951, *d* = −0.005, 95% CI = [−0.15, 0.14]. This measure tests a downstream consequence of habit formation—perceived automaticity. However, perception of automaticity is not an especially sensitive measure in and of itself and could tap processes other than habit (Hagger et al., 2015; Mazar & Wood, 2018). Given that even participants in the weak-habit condition reported relatively high levels of experienced automaticity, our simple key-pressing task seems to have produced uniformly high subjective automaticity. For this reason, we do not discuss this measure further.

Habit strength: reaction time. As in prior research, the strength of habit associations in the key-pressing task was assessed directly through reaction times to respond to the cue (Hardwick et al., 2019). Because of the skewness common in reaction time distributions, median rather than mean reaction times were used in all analyses.

Comprehension check. To ensure that participants understood the help request, we asked them to answer the following prompt immediately after responding: "What was the request that you just responded to?" The options were (a) "To continue for an additional 20 minutes," (b) "To continue for an additional 5 minutes," (c) "To recommend the study to a friend," or (d) "To receive double compensation for my participation." Answers (a), (c), and (d) were coded as incorrect.

Results

Descriptive statistics and correlations among key variables are presented in Table 1. The percentage of participants in each condition who agreed to help is presented in Table 2. Regression models were fitted to the data using the following predictors: mood condition (dummy coded: control condition as the reference level), habit condition (effects coded: −1 = weak habit, +1 = strong habit), and interactions between mood and habit. This analytic design was used for all Study 1 analyses.

Manipulation checks: mood. The mood induction successfully invoked positive affect (measured via the PANAS): In the sad-memory condition, participants reported less positive affect compared with the control condition, *b* = −0.37, 95% CI = [−0.64, −0.10], β = −0.24, *p* = .007. In the happy-memory condition, participants reported more positive affect compared with the control condition, *b* = 0.42, 95% CI = [0.15, 0.69], β = 0.28, *p* = .002. Habit condition did not significantly influence positive affect, *b* = 0.09, 95% CI = [−0.28, 0.11], β = −0.06, *p* = .384. No interactions emerged between habit and mood conditions, both *p*s > .2.

Results for negative affect also indicated the success of the mood induction: Participants in the sad-memory condition reported higher levels of negative affect compared with participants in the control condition, *b* = 0.82, 95% CI = [0.63, 1.02], β = 0.72, *p* < .001. Participants in the happy-memory condition did not differ on negative affect from participants in the control condition, *b* = 0.03, 95% CI = [−0.17, 0.22], β = 0.02, *p* = .793. This is in line with previous work showing the relative independence of the positive and negative PANAS subscales, reflecting an underlying separation of positive and negative affect (e.g., Thompson, 2007). Habit condition did not significantly influence negative affect, *b* < 0.01, 95% CI = [−0.14, 0.15], β < 0.01, *p* = .944. No interactions were found between habit and mood conditions, both *p*s > .2.

As an additional mood-manipulation check, two coders unaware of condition and hypotheses coded each open-ended text response to the mood manipulation for the presence of negative affect (97% agreement, Cohen's κ = .95) and positive affect (95% agreement, Cohen's κ = .91). Disagreements between coders were resolved by discussion. Most open-ended responses in the negative-affect condition showed negative affect (84%), compared with few of the responses in the control (12%) and positive-affect (4%) conditions, suggesting that the mood manipulation was successful. Similarly, most responses in the positive-affect condition showed positive affect (80%) compared with fewer in the control (35%) and negative-affect (4%) conditions.

Manipulation check: habit. The reaction time measure revealed that the habit manipulation was successful: In the strong-habit condition, participants were significantly faster to respond than in the weak-habit condition, *b* = −20.71, 95% CI = [−28.37, −13.05], β = −0.34, *p* < .001. No other effects attained significance (all *p*s > .05).

Table 1. Descriptive Statistics and Correlations for Key Variables in Study 1

Variable	M (SD)	Correlations					
		1	2	3	4	5	6
1. Habit attribution	11.49 (13.90)	—					
2. Mood attribution	28.52 (16.11)	.04 [−.03, .11]	—				
3. Incentivized habit other-attribution	21.33 (13.62)	.49** [.43, .54]	.12** [.05, .19]	—			
4. Incentivized mood other-attribution	32.01 (12.84)	.04 [−.03, .11]	.57** [.52, .61]	.09** [.02, .16]	—		
5. Positive affect	5.26 (1.52)	.05 [−.02, .12]	.11** [.05, .18]	.04 [−.02, .11]	.16** [.09, .23]	—	
6. Negative affect	2.13 (1.15)	.13** [.06, .19]	.07 [−.00, .13]	.11** [.04, .18]	.02 [−.05, .09]	−.20** [−.26, −.13]	—
7. Reaction time	298.8 (61.44)	.05 [−.02, .12]	.08* [.01, .15]	.01 [−.06, .08]	.06 [−.01, .13]	.15** [.08, .22]	−.04 [−.11, .03]

Note: Attributions ranged from 0 to 50 (higher scores reflect greater importance), and mood scores ranged from 1 to 9 (higher numbers reflect stronger feelings). Higher reaction times (given in milliseconds) reflect slower responding in the habit-training task.
*p < .05. **p < .01.

Table 2. Percentage of Participants Who Agreed to Help, by Mood and Habit Condition in Study 1

| Mood condition | Habit condition | | Total |
	Weak	Strong	
Control	60% (50/83)	43% (76/177)	48% (126/260)
Sad	48% (46/96)	47% (87/185)	47% (133/281)
Happy	55% (50/91)	47% (82/176)	49% (132/267)
Total	54% (146/270)	46% (245/538)	48% (391/808)

Note: Higher percentages reflect more participants agreeing to help by working 5 min extra. Raw numbers are given in parentheses.

Habit and mood effects on helping: actual. A logistic regression model tested the actual effects of habit and mood on helping behavior (yes/no). Habit significantly influenced this decision; specifically, participants in the strong-habit condition (who had extensively practiced the "no" response key) were less likely to agree to help, odds ratio (OR) = 0.84, 95% credible interval (CrI) = [0.73, 0.98], p = .022. Mood condition did not significantly influence helping, either for the sad-mood condition, OR = 0.95, 95% CrI = [0.68, 1.33], p = .758, or the happy-mood condition, OR = 1.03, 95% CrI = [0.73, 1.45], p = .855. Thus, habit influenced behavior, whereas mood did not.[1]

Habit and mood effects on helping: attributed. To test the perceived effects of habit and mood, we used a dependent-samples t test to assess the within-participants difference in attributions to mood compared with habit. Participants strongly attributed their behavior to mood over habit (mean difference = 17.07, 95% CI = [15.63, 18.51]), $t(799)$ = 23.22, p < .001, d = 1.14, 95% CI = [1.01, 1.26]. As anticipated, a strong (albeit somewhat smaller) bias in favor of mood remained when participants were incentivized to give accurate attributions for others' behavior (mean difference = 10.66, 95% CI = [9.42, 11.90]), p < .001, d = 0.80, 95% CI = [0.69, 0.91]. Thus, as anticipated, participants' attributions favored mood over habit more than would be expected given the actual effects of each on behavior. It should be noted that despite the strong favoring of mood, habits were judged a plausible explanation, especially for other people's behavior: Habits received an importance rating of 21% (out of 50%), compared with 32% for mood.

Exploratory analyses: intensity of experience and attributions. We explored whether participants with stronger moods and stronger habits were more likely to make attributions to mood and habit, respectively. In general, participants with stronger internal states were only slightly more likely to attribute their behavior to these states. That is, attribution to mood was weakly correlated with positive PANAS scores, $r(801)$ = .11, and marginally correlated with negative PANAS scores, $r(801)$ = .07 (see Table 1). In addition, attribution to mood was weakly correlated with reaction time on the habit-formation task, so that participants with slower reaction times gave stronger mood attributions, $r(801)$ = .08. Attributions to habit showed only a slight positive correlation with negative affect, $r(800)$ = .13. Thus, the attributional bias we documented was not limited to people with intense moods or weak habits.

Discussion

This first study provided causal evidence that people's explanations for their behavior favor inner states over habits, even when that behavior is driven by habit. We manipulated habit strength via amount of practice at a key-press task and manipulated mood through a memory-recall task. Participants then indicated their willingness to help by pressing a highly practiced or less-practiced key.

In our test of actual influences on helping, habit strength determined helping, but current mood did not. Specifically, participants in the strong-habit condition, who had earlier extensively practiced the "no" response key, were more likely to decline a helping request compared with participants in the weak-habit condition, who had practiced the "yes" and "no" response keys equally. Thus, participants' decisions continued to be influenced by their prior key-pressing habits. In contrast, participants induced to feel sad or happy helped at rates comparable with those of participants in a control condition who experienced no mood manipulation. Note that our hypotheses were not about these

behavioral effects per se but instead concerned the difference between actual and perceived effects of habit and mood on behavior.

When explaining their behavior, participants attributed their helping more to current mood than to habit. Thus, their attributions were misaligned with the actual determinants of behavior—they underestimated habit and overemphasized mood. Our design provided a compelling test of this hypothesis, given that participants in the strong-habit condition should be aware of their recent, extensive practice at pressing a particular computer key. However, when incentivized to make accurate attributions about other people's behavior, participants still revealed a substantial attribution gap favoring mood over habit. Thus, it does not seem that the attribution pattern was due to artifacts of social desirability or to conversational norms that favor mood explanations. The incentivized measure also suggested that attributions to habit were meaningful: Despite the substantial attribution gap favoring mood over habit, participants considered habit to be a plausible determinant of other people's behavior.

Study 2

Study 2 investigated the attribution bias with a mundane repeated behavior—coffee drinking—recorded over the course of a typical week. Specifically, coffee drinking was assessed in response to an inner state (fatigue) and to habit strength, which are common reasons for coffee drinking (see Pilot Study below). The second study thus moved beyond Study 1, which tested explanations for a single behavior immediately following an emotionally evocative experience designed to enhance the salience of mood.

Given that some of our participants drank coffee very often, we anticipated that habit (more than fatigue) would strongly influence actual coffee drinking. However, as in the first study, we anticipated that participants' attributions would emphasize fatigue as much as or more than habit. Thus, these two hypotheses together concern the correspondence between the actual and perceived determinants of behavior. In an additional test of our model, we anticipated that fatigue attributions would be unrelated to within-participants associations between fatigue and coffee drinking. In other words, people's beliefs about fatigue determining their coffee drinking would be unrelated to its actual role in driving their individual consumption.

Method

Pilot study. To assess lay beliefs about the causes of coffee drinking, we asked 40 college students (22 male,

16 female, 2 genderqueer or other) to rate six causes of coffee drinking on 5-point scales (1 = *not at all important* to 5 = *extremely important*). These causes were fatigue ("tiredness or low energy"), habit ("habit or behavior routines"), thirst, taste, social motives ("spending time with friends"), and coffee after a meal. Fatigue was rated as most important ($M = 4.05$, $SD = 0.96$), followed by taste ($M = 3.58$, $SD = 1.03$), habit ($M = 3.50$, $SD = 1.22$), social motives ($M = 3.17$, $SD = 1.08$), having coffee after a meal ($M = 2.12$, $SD = 1.22$), and thirst ($M = 1.70$, $SD = 0.91$). A paired-samples t test comparing fatigue and habit attributions (within participants) revealed that participants attributed coffee drinking to fatigue significantly more than to habit, mean difference = 0.55, 95% CI = [0.14, 0.96], $t(39) = 2.72$, $p = .01$, $d = 0.50$, 95% CI = [0.11, 0.88].

Design. To capture experiences and explanations as they naturally unfold in daily life, we used a combination of surveys, daily morning reports, and EMA. Participants first completed intake surveys, including measures of habit strength and attributions for their own coffee drinking. Then, over the course of a week, they reported every 2 hr on their fatigue and coffee drinking. They also completed a brief survey every morning immediately after waking up.

Our analysis predicted coffee drinking at one prompt from fatigue experienced at the prior prompt. This lagged design minimized any self-report bias that might emerge from concurrent associations between fatigue and coffee drinking (i.e., "I'm drinking coffee therefore I must be tired"). After the study week, participants completed a final survey. Finally, participants completed a follow-up survey when data collection for the study ended.

Power analyses. EMA designs such as the present one typically generate thousands of prompts (Level 1 sample size), which tend to produce very high power for within-participants effects. Because most of our research questions could be probed within participants (e.g., using our novel context-specific habit measure), we aimed for a final sample size of 120, which is in line with typical sample sizes in EMA studies (cf. a mean sample size of 99 in a recent systematic review; Wen et al., 2017).

To estimate observed power for our multilevel logistic regression, we simulated a data set with log-odds regression coefficients of 0.3 and 0.2 (corresponding to our *OR*s of 1.35 and 1.22) for a Level 2 and Level 1 variable, respectively. Simulated sampling from this data set 1,000 times revealed that 50 participants were sufficient to achieve 90% power for our between-participants variable (habit strength) and 99.5% power for our within-participants variable (fatigue).

Participants. Participants were a convenience sample of 112 U.S. undergraduate students who received either course credit or monetary compensation (27 male, 85 female; age: *M* = 20.85 years, *SD* = 2.85, range = 18–33). The (self-reported) selection criteria were (a) speaking English fluently, (b) owning a smartphone, (c) being 18 or older, and (d) drinking coffee at least once a week. An additional 35 participants were excluded for drinking coffee once or less often during the study period, and 4 additional participants were excluded for answering fewer than 50% of prompts. Thus, the final sample for analyses was slightly smaller than our preregistered target of 120.

To minimize attrition, we linked compensation to compliance. Paid participants received $20 for completing 80% to 100% of EMA prompts, $15 for completing 50% to 80%, and $5 for completing fewer than 50%. Participants who received course credit had a similar three-tiered compensation system.

Procedure.

Intake session. After providing informed consent, participants reported the strength of their coffee-drinking habit, coffee-drinking intentions and attitudes, coffee-drinking attributions, and demographic information (see the Measures section below). In addition, to obscure the purpose of the study and limit reactivity, we asked participants to answer an identical set of measures about soft drinks. Participants then wrote down implementation intentions (Adriaanse et al., 2011) to overcome potential obstacles for completing the prompts (e.g., "If my phone beeps when I am with people, then I will excuse myself and answer the prompt").

Ecological momentary assessment. For 5 weekdays (participants were not prompted on Saturday and Sunday), participants were prompted to respond eight times per day at regular 2-hr intervals from 8:00 a.m. to 10:00 p.m. Each prompt included items meant to obscure the purpose of the study, including location (e.g., home, campus) and temperature (hot, cold, or comfortable). Participants then reported how tired they were, whether they drank coffee in the past 2 hr, and whether they drank soft drinks in the past 2 hr. Participants also completed an exploratory mood item and an open-ended response item in which they briefly described their current situation.

In addition, because fatigue on waking up may be particularly important for coffee drinking, participants completed a prompt every morning on getting out of bed. Morning prompts included the items in the regular prompts as well as an item asking whether they had already drunk coffee (a measure of compliance). Thus, we could measure the prospective effect of waking fatigue on coffee drinking and avoid the self-report bias that might emerge with concurrent reports.

At the end of the first study day, participants with response rates of 50% or above (four or more prompts) were informed of their approximate level of compliance via email (50%–75% or 75%–100%). Participants with less than 50% compliance were contacted by phone, text message, or both to address potential technical difficulties that might have led to low compliance.

Final survey. Participants were sent the final survey over the weekend after they completed the EMA portion of the study. This survey included the context-specific habit measure, the single-event self-attribution measure, and open-ended text measures asking about self-regulation and general study feedback.

Follow-up survey. Shortly after all data collection was completed, participants were emailed a survey that included the incentivized self-attribution measure (see below).

Measures: intake. Additional measures for this study are included in the Supplemental Material.

Habit strength. Using the Behavior Frequency in Context (BFiC) measure (Galla & Duckworth, 2015; Ji & Wood, 2007), participants reported on a 5-point scale how often they drink coffee (1 = *less than once a week* to 5 = *more than 7 times a week; that is, more than once a day*). They then rated on 5-point scales how often they drink coffee at the same time of day and at the same location (1 = *never or almost never at the same* [*time/location*] to 5 = *almost always or always at the same* [*time/location*]). Each participant's rating of coffee-drinking frequency was then multiplied by the time and location stability ratings separately, and the two Frequency × Context scores were averaged to create a mean habit-strength score.

Participants also completed the SRHI (Verplanken & Orbell, 2003), which is the full version of the brief four-item Self-Reported Behavior Automaticity Index used in Study 1 (note that we did not include the full SRHI in Study 1 to minimize participant burden). Participants indicated their agreement with a set of 11 statements regarding coffee drinking (e.g., "drinking coffee is something that I do without thinking," "drinking coffee is something that belongs to my daily routine") on 7-point scales (1 = *strongly disagree* to 7 = *strongly agree*).

Self-attribution. On percentage scales (0% to 100%), participants rated the extent to which their coffee drinking was driven by habit ("my past behavior and habits")

and by fatigue ("my energy levels and tiredness"). The anchors were 0%, which indicated that coffee drinking was unaffected by a factor, and 100%, which indicated that coffee drinking was completely determined by a factor. Participants were instructed not to allow the sum of both ratings to exceed 100%.

Coffee-drinking intentions and attitudes. On a 100-point scale (0 = *not at all* to 100 = *extremely*), participants rated how much they liked drinking coffee ("how much do you enjoy drinking coffee?"). On a 7-point scale (1 = *strongly disagree* to 7 = *strongly agree*), participants rated their coffee-drinking intentions ("I intend to drink coffee . . . "; the ellipsis was replaced by the individual participant's self-reported frequency of coffee drinking; Ajzen, 2002).

Measures: EMA.

Fatigue. Using a 6-point scale, participants rated how tired they were (1 = *not at all* to 6 = *extremely*).

Coffee drinking. To ensure that a single coffee consumed over a period of time was categorized as one episode, we asked participants to indicate whether they started drinking coffee in the past 2 hr (responses were "No," "Yes - 1 Drink," "Yes - 2 Drinks," "Yes - 3 Drinks or more"). Answer choices were categorized into a binary drink/did-not-drink indicator of coffee drinking.

Mood. In this exploratory measure, participants rated on a 5-point scale their current mood (1 = *unhappy* to 5 = *happy*).

Situation description (open-ended measure). In an open-ended response, participants briefly described their current situation (e.g., "going to the gym," "with friends"). Specifically, for prompts in which they indicated that they had recently been drinking coffee, they described that coffee-drinking situation. For prompts in which they reported not drinking coffee, they described the situation they were in 1 hr previously. These situation descriptions were then used in the context-specific habit measure and the single-event attribution measure (see below).

Measures: final survey.

Context-specific habit measure. As an exploratory measure, seven situation descriptions were randomly selected for each participant to reflect prompts in which they drank coffee and prompts in which they did not drink coffee. For each situation description, participants rated (a) how often they drank coffee in that situation, (b) how automatic they perceived coffee drinking to be in that situation, and (c) the strength of their intentions to drink coffee in that situation (see the Supplemental Material for full descriptions).

Single-event self-attribution. To evaluate attributions for a specific instance of a behavior, we showed participants their open-ended text situation description for their own final coffee-drinking event and asked them to rate the extent to which habit and fatigue contributed to drinking coffee at that time. Item wording and answer choices were the same as in the intake-attribution measure. To confirm that participants recalled the specific coffee-drinking event, we asked whether they remembered it, and analyses for this single-event measure included only the 81 participants who answered affirmatively (72%).

Measures: follow-up survey. Participants were offered a monetary incentive of $3 if they accurately estimated the effects of fatigue and habit strength on their own coffee drinking during the study week. The incentive, along with using their own data as an objective benchmark, was designed to encourage participants to respond accurately and reduce any influences from social desirability or conversational norms. A total of 78 (70%) participants responded to the follow-up survey and thus provided this rating.

Results

Means, standard deviations, and between-person correlations for key variables appear in Table 3. The 112 participants (Level 2 sample size) produced 3,550 individual observations (Level 1 sample size), corresponding to an average response rate of 31.7 out of 40 EMA prompts (79%).

On average, participants drank coffee a little over five times during the 5-day period (*M* = 5.26, *SD* = 3.09), or approximately once a day. Scores on both habit-strength measures suggested moderate coffee-drinking habits. Furthermore, the two measures were strongly correlated with each other, $r(109) = .73$, 95% CI = [.63, .81]. Choice of habit measure did not have a noticeable impact on the results, and thus analyses are reported using the BFiC scale (see Table S2 in the Supplemental Material for analysis results using the SRHI).

Primary analyses. Results were analyzed using the following multilevel model:

Level 1: $\text{coffee}_{ij} = \beta_{0j} + \beta_{1j}\text{fatigue\_cmc}_{ij} + e_{ij}$

Level 2:

$$\beta_{0j} = \gamma_{00} + \gamma_{01}\text{habit}_j + \gamma_{02}\text{mean\_fatigue}_j$$
$$+ \gamma_{03}\text{attribution}_j + u_{0j}$$

$$\beta_{1j} = \gamma_{10} + \gamma_{11}\text{habit}_j + \gamma_{12}\text{attribution}_j + u_{1j}.$$

Table 3. Descriptive Statistics and Between-Participants Correlations for Key Variables in Study 2

		Correlations					
Variable	M (SD)	1	2	3	4	5	6
1. Habit strength (BFiC)	11.79 (5.21)	—					
2. Habit strength (SRHI)	4.09 (1.35)	.73** [.63, .81]	—				
3. Mean fatigue	3.44 (0.75)	-.05 [-.23, .14]	.09 [-.09, .28]	—			
4. Habit attribution	36.99 (27.59)	.41** [.25, .56]	.46** [.30, .60]	-.07 [-.25, .12]	—		
5. Fatigue attribution	69.14 (23.43)	-.34** [-.49, -.16]	-.24* [-.41, -.05]	-.03 [-.22, .16]	-.37** [-.52, -.20]	—	
6. Coffee count[a]	5.26 (3.09)	.54** [.39, .66]	.44** [.28, .58]	-.08 [-.26, .11]	.20* [.01, .37]	-.14 [-.32, .04]	—
7. Response rate	31.70 (4.42)	-.05 [-.23, .14]	-.06 [-.24, .13]	-.01 [-.19, .18]	-.16 [-.33, .03]	-.13 [-.31, .06]	.03 [-.16, .21]

Note: Scores for the Behavior Frequency in Context (BFiC) scale ranged from 1 to 25, and scores for the Self-Report Habit Index (SRHI) ranged from 1 to 7; higher numbers reflect stronger habits. Mean fatigue ranged from 1 to 6; higher numbers reflect higher fatigue. Habit attribution and fatigue attribution ranged from 1 to 100; higher scores reflect stronger attributions. Response rates reflect the number of prompts answered (out of 40 possible); higher scores indicate higher response rates. Values in brackets are 95% confidence intervals.
[a]Coffee count is the total number of coffee-drinking events reported during the study period.
*p < .05. **p < .01.

Table 4. Coefficient Estimates for Fixed Effects in the Multilevel Model Predicting Coffee Drinking in Study 2

Variable	Odds ratio	95% credible interval
Habit strength, γ_{01}	1.35	[1.16, 1.55]
Fatigue (person mean centered), γ_{00}	1.22	[1.08, 1.39]
Fatigue (average person level), γ_{02}	0.96	[0.84, 1.09]
Fatigue attribution, γ_{03}	1.02	[0.89, 1.18]
Habit Strength × Fatigue (person mean centered) interaction, γ_{11}	0.95	[0.84, 1.07]
Fatigue (person mean centered) × Fatigue Attribution, γ_{12}	0.96	[0.86, 1.09]

Note: The 95% credible interval represents the range of values that has a 95% chance of including the population odds ratio.

where *i* and *j* represent observations (*i*) nested within participants (*j*). *Coffee* refers to whether the participant did or did not report drinking coffee in the following prompt (i.e., a lead indicator of coffee drinking, meant to capture the prospective association between fatigue and coffee drinking in the following 2 hr). This lagged design controls for response biases associated with concurrently reporting a predictor and outcome. *Fatigue_cmc* is a person-mean-centered fatigue rating at each EMA prompt, computed by subtracting each participant's mean level of fatigue from each fatigue rating.

Positive values reflect higher-than-average fatigue for that person, and negative values reflect lower-than-average fatigue for that person. *Mean_fatigue* is each person's mean level of fatigue, *habit* is each participant's habit-strength score, and *attribution* is each person's attribution of coffee drinking to fatigue.

Habit and fatigue effects on coffee drinking: actual.
Model estimates for the primary multilevel model are shown in Table 4 and Figure 1. To facilitate interpretation and reduce multicollinearity, we standardized all predictors
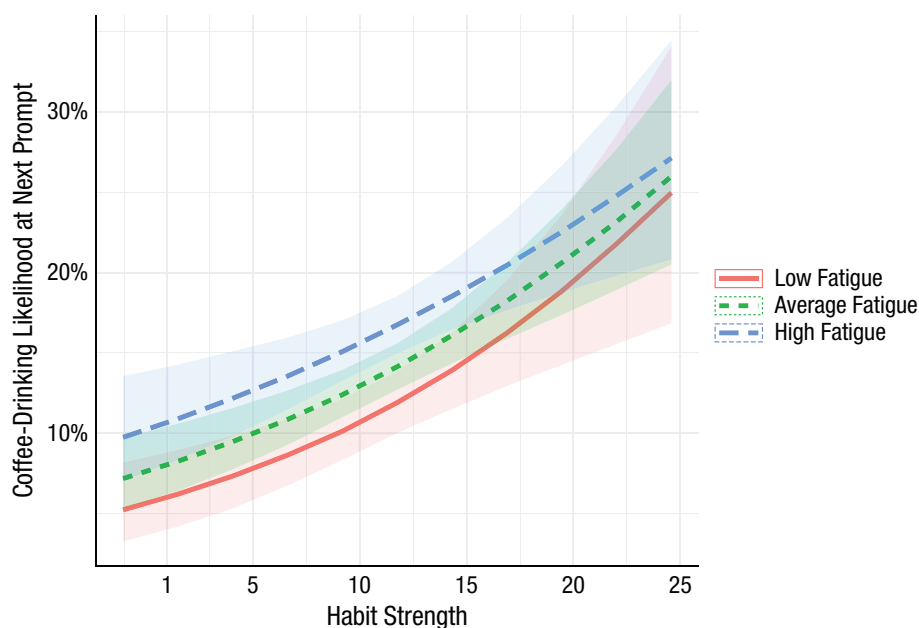

Fig. 1. Likelihood of having drunk coffee by the time of the ecological-momentary-assessment prompt as a function of the strength of participants' coffee-drinking habit and the amount of fatigue reported at the prior prompt (Study 2). For fatigue scores, low and high refer to −1 *SD* and +1 *SD*, respectively, from each participant's average. Error bands indicate 95% confidence intervals.

in all regression analyses below to have a mean of 0 and standard deviation of 1. Because of convergence issues with the original frequentist model, we respecified the main model as Bayesian. To avoid imposing restrictive priors on the results, we specified an uninformative prior for all model predictors (a prior slope value of 0 with a standard deviation of 100).

To test whether habit determined coffee drinking as well as or better than fatigue, we compared the standardized coefficients for habit strength and person-mean-centered fatigue (γ_{01} and γ_{10} in the model). As anticipated, participants with stronger habits were more likely to drink coffee, $OR = 1.35$, 95% CrI = [1.16, 1.55]. Yet participants also drank more coffee after being fatigued (within participants), $OR = 1.22$, 95% CrI = [1.08, 1.39]. CrIs for all other model effects spanned 1.00 (see Table 4).

To determine whether these effects also held for the first coffee of the day, we computed a separate multilevel analysis that predicted coffee drinking on the first scheduled prompt of each day (i.e., excluding the participant-initiated morning prompts) from waking fatigue as measured by morning prompts, habit strength, and an interaction between habit and waking fatigue. The final sample size for this analysis consisted of 102 participants (Level 2 sample size) and 330 responses (Level 1 sample size, corresponding to participant days). Out of the original 498 responses, 46 were excluded because participants reported that they had already drunk coffee by the time they completed the prompt; an additional 122 morning reports were excluded because they were submitted after the first EMA prompt of that day and therefore could not be used to predict drinking in that prompt. The analysis revealed that early-morning fatigue was unrelated to coffee drinking by the following prompt, $OR = 0.99$, 95% CrI = [0.67, 1.45], $p = .950$. Also, participants with stronger habits were more likely to drink first thing in the morning, $OR = 2.18$, 95% CrI = [1.38, 3.44], $p = .001$. The interaction between habit strength and fatigue was not significant, $p > .2$. Thus, consistent with our expectations, results showed that waking fatigue did not influence coffee drinking on the first prompt of the day, whereas habit strength did.

Given that the present study measured rather than manipulated habit, we examined whether habit uniquely determined coffee drinking over and above the contribution of attitudes or intentions concerning coffee drinking. When these two predictors were added to the main model, the results remained essentially unchanged: There were significant main effects of habit strength and fatigue but no effects for liking or intentions (see Table S3 in the Supplemental Material). Thus, consistent

with our hypotheses, results showed that the effect of habit strength on coffee drinking was not due to liking for coffee or intentions to drink it.

Habit and fatigue effects on coffee drinking: attributed. To test our hypothesis that participants would attribute coffee drinking to fatigue more than habit, we computed a paired-samples t test comparing the within-participants difference between each participant's fatigue and habit attributions. As expected, fatigue attributions were significantly stronger than habit attributions, mean difference = 32.14, 95% CI = [24.18, 40.11], $t(110) = 8.00$, $p < .001$, $d = 1.26$, 95% CI = [0.84, 1.67].

To ensure that the attribution results were not due to a failure to recall coffee-drinking events or to ambiguity of attributions for multiple instances of a behavior, we evaluated the single-event attribution measure for participants' last coffee-drinking episode. Consistent with our hypotheses, results of a paired-samples t test revealed that, just as with attributions for overall coffee drinking, participants attributed their most recent coffee-drinking event to fatigue more than habit, mean difference = 25.86, 95% CI = [15.65, 36.08], $t(80) = 5.04$, $p < .001$, $d = 0.94$, 95% CI = [0.50, 1.38].

Suggesting that the attribution findings are not due to social desirability or conversational norms, results of a paired-samples t test with the incentivized self-attribution measure, designed to maximize accuracy, revealed significantly stronger fatigue attributions than habit attributions, mean difference = 16.01, 95% CI = [5.82, 26.20], $t(77) = 3.13$, $p = .002$, $d = 0.62$, 95% CI = [0.19, 1.04]. Thus, incentivized participants still overwhelmingly rated fatigue as more important than habit, even though incentives reduced the size of this effect (a mean difference of about 16 on the 100-point scale when incentivized, compared with about 32 in the nonincentivized measure).

Alternative habit measure: effects of context-specific habit: actual. The final survey assessed a novel, within-participants measure of habit strength to directly compare with our within-participants measure of fatigue. A multilevel model predicted actual coffee drinking at each prompt (yes/no) from fatigue (within participants) and context-specific habit (within participants). As anticipated, context-specific habit was an especially strong determinant of coffee drinking, $OR = 1.86$, 95% CrI = [1.52, 2.32]. As in the main analysis, fatigue predicted coffee drinking as well, $OR = 1.24$, 95% CrI = [1.05, 1.49]. The larger effect of habit compared with fatigue, as well as the nonoverlapping credible intervals, reveal that this measure of habit exerted a stronger effect on coffee drinking than fatigue.

Correspondence between perceived and actual effects. This study's repeated longitudinal design allowed us to estimate not only the overall effect of fatigue but also whether people who attributed coffee drinking to fatigue were the ones who actually drank more coffee when fatigued. If participants' attributions to fatigue were based on shared cultural theories (e.g., Wilson et al., 1982) rather than personal experience, however, we expected to find that attributions to fatigue were unrelated to the actual within-participants effect of fatigue on coffee drinking. For this analysis, our main multilevel model tested whether attribution to fatigue moderated the within-participants lagged association between fatigue and coffee drinking (γ_{12} in the model). A positive slope would suggest that participants with stronger associations between fatigue and coffee drinking also gave stronger fatigue attributions (i.e., more accurate attributions). However, supporting the cultural-theory account, results showed that strength of attribution to fatigue did not moderate the association between within-participants fatigue and coffee drinking at the next prompt, $OR = 0.95$, 95% CrI = [0.82, 1.11]. Thus, participants strongly attributed coffee drinking to fatigue regardless of fatigue's actual effect on their own coffee drinking, consistent with the notion that attributions draw on shared lay theories.

Exploratory analyses: intensity of experience and attributions. Correlational analyses provided additional insight into the accuracy of participants' attributions (see Table 3). First, attributions to fatigue were not correlated with mean fatigue levels, $r(109) = -.03$. This lack of effect is consistent with the weak correlations in Study 1 between mood intensity and attributions, along with prior findings that attributions often reflect shared cultural lay theories more than individual experience (Wilson et al., 1982).

Second, on a correlational basis, participants with stronger coffee habits made stronger habit attributions, $r(109) = .41$ (behavior frequency in context), $r(108) = .46$ (self-report habit index). To probe this effect, we divided our sample into tertiles by habit strength (weak, moderate, and strong). To enable us to compare the attribution measure with the actual influences on coffee drinking, we computed our main multilevel model predicting coffee drinking by extracting within-participants effects for fatigue and habit for each participant (using the context-specific habit measure; to match the attribution measure, model slopes in log-odds units were standardized to have a range of 0–100%). Participants in the weak-habit condition attributed on average a 51% difference in favor of fatigue, compared with an actual difference of 18% in the opposite direction (i.e., in favor of habit). Participants with moderate habits attributed a 39% difference in favor of fatigue, compared with a 21% difference in favor of habit. Participants in the strong-habit condition attributed 4% in favor of fatigue, compared with an actual difference of 33% in favor of habit. Thus, participants with stronger habits correctly made stronger habit attributions and weaker fatigue attributions, but they continued to favor fatigue more than was merited by the actual predictors of their behavior.

Exploratory analyses: downstream effects of attribution accuracy on well-being. To identify downstream effects of attributions, we assessed whether attribution to habit over fatigue is associated with a more positive mood in life, as measured using the average of participants' mood reports in the EMA prompts. Attribution scores were calculated as the difference between each participant's attributions to mood and to habit; positive scores implied attribution to fatigue more than habit, and negative scores implied attribution to habit more than fatigue. Attribution scores were moderately and negatively correlated with mood, $r(109) = -0.27$, 95% CI = [−0.44, −0.09], so that attribution to habit over fatigue was associated with more positive mood.

To explore whether attributions favoring habit exert a unique effect on mood or whether this correlation is simply due to people who are less tired and have stronger habits being more happy, we fitted a linear regression model predicting mood from participants' habit strength (measured using the BFiC scale), mean fatigue levels, and attribution difference scores. Supporting a unique effect of attribution, results showed that stronger attribution to habit over fatigue predicted more positive mood, $b = -0.12$, 95% CI = [−0.23, −0.02], $\beta = -0.23$. Lower mean fatigue levels were also associated with more positive mood, $b = -0.22$, 95% CI = [−0.31, −0.13], $\beta = -0.40$. Habit strength did not show a discernible association with mood, $b = 0.04$, 95% CI = [−0.06, 0.14], $\beta = 0.07$. Thus, these exploratory analyses suggest that greater accurate recognition of habit in one's own behavior is associated with higher well-being.

Discussion

In this second study, participants explained the causes of a mundane everyday action—coffee drinking—and then tracked their momentary fatigue and coffee drinking over the course of a typical week. Again, our hypotheses compared the actual influences on behavior with participants' behavioral attributions. Fatigue and habit strength had comparable effects on actual behavior. The strong effect of habit held across three different

measures of habit. Furthermore, analyses of the first coffee drink of the day and the within-participants habit measure supported our hypothesis that participants would drink in response to habit more than fatigue. Thus, if participants' attributions were accurate, they should have featured habit as much or more than fatigue. However, participants miscalibrated these behavioral influences by attributing their coffee drinking primarily to fatigue rather than habit.

Notably, Study 2 revealed a bias to overlook habit despite design features included to reduce misattribution. The inaccuracy in self-attribution persisted when participants were incentivized for accuracy and when asked at the end of the study about a recent coffee-drinking event rather than their coffee drinking in general. Thus, inaccurate attributions emerged despite motivation, an objective criterion, and adequate opportunity to observe their behavior given the frequency of consumption (once a day on average). Further attesting to the robustness of this attribution bias, coffee drinking is a mundane, everyday action that, unlike the task in our first study, is not commonly preceded by a salient emotion-inducing experience.

General Discussion

In two studies, participants' explanations overemphasized inner states and underemphasized habit. Participants' actual willingness to donate time in a laboratory task as well as their everyday coffee drinking were determined as much or more by habits than by inner states (mood and fatigue, respectively). However, participants' attributions for why they acted the way they did emphasized inner states more than habit. Thus, participants appeared to be both undervaluing habit compared with its actual influence on behavior and overvaluing inner states such as mood and fatigue. This pattern is understandable given the disproportionate value that people place on personal introspections (Pronin, 2009) as well as general cognitive and motivational tendencies to interpret actions as goal directed (Rosset, 2008). Through these forces, people may form socially shared lay theories about behavior that inform their attributions. This *lure of phenomenology* not only biases lay theories but also may have oriented formal psychological theories to overvalue salient, motivational determinants of behavior (see Duckworth et al., 2016).

The combination of experimental manipulation in Study 1 and naturalistic observation in Study 2 provides evidence for the causal role of habits as well as the relevance of this attribution bias in everyday settings. Furthermore, the results were replicated across the different measures of habit strength appropriate in these different tasks: Study 1's manipulation of practice along with a reaction time measure and Study 2's self-report measures of behavioral repetition in a given context (a determinant of habit formation), experienced automaticity (a consequence of habit formation), and an exploratory within-participants, context-specific habit measure.

A number of features of our research would be expected to maximize participants' accuracy and minimize biases. Each study assessed attributions for specific recent behaviors, minimizing the opportunity for biased recall. Furthermore, in both studies, attributional biases were evident even when participants were incentivized for accuracy, as well as when participants were explaining others' behavior (Study 1) and regardless of whether participants were explaining a repeated behavior in general or in a specific recent instance (Study 2).

That our participants discounted habit's influences on their own behavior may seem at odds with the *actor-observer effect*, which suggests that people have a bias to attribute their own behavior more to environmental factors than they do others' behavior (Jones & Nisbett, 1971). Yet a meta-analysis of the literature showed that the actor-observer effect on attributions emerges largely under specific conditions, such as for negative rather than positive events (Malle, 2006). Furthermore, habits are not solely an environmental influence, as they reside both in memory and in the environment that triggers a habitual response.

Both studies recruited fluent English speakers, predominantly residing in the United States and the United Kingdom. If habit underestimation depends on agency beliefs, it may be smaller in collectivistic cultures that place less emphasis on individuals and more emphasis on context (e.g., Crandall et al., 2001). Furthermore, because the recruited samples included online panel participants (in Study 1) and predominantly female college students (Study 2), it is unknown to what extent the findings generalize to other populations.

Similar to other lay-theory biases, habit underestimation may give rise to important downstream effects. For example, it raises questions about the accuracy of people's reports that lack of willpower is a primary reason for personal failure to lose weight, save money, and exercise (American Psychological Association, 2012). It also raises questions about the effectiveness of common self-regulation strategies: If people misattribute the sources of their behavior, then they may focus on strategies that affect inner states (e.g., limit coffee drinking by reducing fatigue) at the expense of situational strategies that may more successfully modify habits. This would align with the argument that situational self-regulation strategies are low in salience, which could lead people to overlook these interventions' potential (Duckworth et al., 2016). Exploratory analyses in Study

2 revealed that participants who placed more weight on habit in their attributions also reported more positive mood, which suggests that accurate attributions are beneficial. It may be that well-being increases not only with habit performance (Heintzelman & King, 2019), but also with recognizing habits' elusive yet pervasive role in daily life.

Transparency

ORCID iD

Asaf Mazar   https://orcid.org/0000-0003-1459-0542

Acknowledgments

Supplemental Material

Additional supporting information can be found at http://journals.sagepub.com/doi/suppl/10.1177/09567976211045345

Note

1. We also ran an exploratory model that included the Habit × Mood interaction, which essentially replicated the results reported in the text: the hypothesized effect of habit, *OR* = 0.71, 95% CrI = [0.54, 0.92], *p* = .011; no effect of the sad-mood condition, *OR* = 0.84, 95% CrI = [0.59, 1.21], *p* = .348, or happy-mood condition, *OR* = 0.96, 95% CrI = [0.66, 1.39], *p* = .826; and only an unexpected nonsignificant interaction between habit and the happy-mood condition, *OR* = 1.38, 95% CrI = [0.96, 1.99], *p* = .078.

References

Adriaanse, M. A., Vinkers, C. D. W., De Ridder, D. T. D., Hox, J. J., & De Wit, J. B. F. (2011). Do implementation intentions help to eat a healthy diet? A systematic review and meta-analysis of the empirical evidence. *Appetite*, *56*, 183–193. https://doi.org/10.1016/j.appet.2010.10.012

Ajzen, I. (2002). *Constructing a TpB Questionnaire: Conceptual and methodological considerations*. http://chuang.epage.au.edu.tw/ezfiles/168/1168/attach/20/pta_41176_7688352_57138.pdf

American Psychological Association. (2012). *Stress in America: Our health at risk*. https://www.apa.org/news/press/releases/stress/2011/final-2011.pdf

Avni-Babad, D. (2011). Routine and feelings of safety, confidence, and well-being. *British Journal of Psychology*, *102*(2), 223–244. https://doi.org/10.1348/000712610X513617

Crandall, C. S., D'Anello, S., Sakalli, N., Lazarus, E., Nejtardt, G. W., & Feather, N. T. (2001). An attribution-value model of prejudice: Anti-fat attitudes in six nations. *Personality and Social Psychology Bulletin*, *27*(1), 30–37. https://doi.org/10.1177/0146167201271003

Duckworth, A. L., Gendler, T. S., & Gross, J. J. (2016). Situational strategies for self-control. *Perspectives on Psychological Science*, *11*, 35–55. https://doi.org/10.1177/1745691615623247

Galla, B. M., & Duckworth, A. L. (2015). More than resisting temptation: Beneficial habits mediate the relationship between self-control and positive life outcomes. *Journal of Personality and Social Psychology*, *109*(3), 508–525. https://doi.org/10.1037/pspp0000026

Gardner, B., Abraham, C., Lally, P., & de Bruijn, G. J. (2012). Towards parsimony in habit measurement: Testing the convergent and predictive validity of an automaticity subscale of the Self-Report Habit Index. *International Journal of Behavioral Nutrition and Physical Activity*, *9*, Article 102. https://doi.org/10.1186/1479-5868-9-102

Gershman, S. J., Gerstenberg, T., Baker, C. L., & Cushman, F. A. (2016). Plans, habits, and theory of mind. *PLOS ONE*, *11*(9), Article e0162246. https://doi.org/10.1371/journal.pone.0162246

Hagger, M. S., Rebar, A. L., Mullan, B., Lipp, O. V., & Chatzisarantis, N. L. D. (2015). The subjective experience of habit captured by self-report indexes may lead to inaccuracies in the measurement of habitual action. *Health Psychology Review*, *9*(3), 296–302. https://doi.org/10.1080/17437199.2014.959728

Hardwick, R. M., Forrence, A. D., Krakauer, J. W., & Haith, A. M. (2019). Time-dependent competition between goal-directed and habitual response preparation. *Nature Human Behaviour*, *3*(12), 1252–1262. https://doi.org/10.1038/s41562-019-0725-0

Heintzelman, S. J., & King, L. A. (2019). Routines and meaning in life. *Personality and Social Psychology Bulletin*, *45*(5), 688–699. https://doi.org/10.1177/0146167218795133

Ji, M. F., & Wood, W. (2007). Purchase and consumption habits: Not necessarily what you intend. *Journal of Consumer Psychology*, *17*, 261–276. https://doi.org/10.1016/S1057-7408(07)70037-2

Jones, E. E., & Nisbett, R. E. (1971). The actor and the observer: Divergent perceptions of the causes of behavior. In E. E. Jones, D. E. Kanouse, H. H. Kelly, R. E. Nisbett, S. Valins, & B. Weiner (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 79–94). General Learning Press.

Knowlton, B. J., & Diedrichsen, J. (2018). Editorial overview: Habits and skills. *Current Opinion in Behavioral Sciences*, *20*, iv–vi. https://doi.org/10.1016/j.cobeha.2018.02.009

Malle, B. F. (2006). The actor-observer asymmetry in attribution: A (surprising) meta-analysis. *Psychological Bulletin*, *132*(6), 895–919. https://doi.org/10.1037/0033-2909.132.6.895

Mazar, A., & Wood, W. (2018). Defining habit in psychology. In B. Verplanken (Ed.), *The psychology of habit: Theory, mechanisms, change, and contexts* (pp. 13–29). Springer. https://doi.org/10.1007/978-3-319-97529-0_2

McFerran, B., & Mukhopadhyay, A. (2013). Lay theories of obesity predict actual body mass. *Psychological Science*, *24*(8), 1428–1436. https://doi.org/10.1177/0956797612473121

Mills, C., & D'Mello, S. (2014). On the validity of the autobiographical emotional memory task for emotion induction. *PLOS ONE*, *9*(4), Article e95837. https://doi.org/10.1371/journal.pone.0095837

Pronin, E. (2009). The introspection illusion. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 41, pp. 1–67). Academic Press. https://doi.org/10.1016/S0065-2601(08)00401-2

R Core Team (2021). *R: A language and environment for statistical computing* (Version 4.1.0) [Computer software]. https://www.R-project.org/

Rosset, E. (2008). It's no accident: Our bias for intentional explanations. *Cognition*, *108*(3), 771–780. https://doi.org/10.1016/j.cognition.2008.07.001

Shiffman, S., Hufford, M., Hickcox, M., Paty, J. A., Gnys, M., & Kassel, J. D. (1997). Remember that? A comparison of real-time versus retrospective recall of smoking lapses. *Journal of Consulting and Clinical Psychology*, *65*(2), 292–300. https://doi.org/10.1037/0022-006X.65.2.292.a

Thompson, E. R. (2007). Development and validation of an internationally reliable short-form of the Positive and Negative Affect Schedule (PANAS). *Journal of Cross-Cultural Psychology*, *38*(2), 227–242. https://doi.org/10.1177/0022022106297301

Verplanken, B., & Orbell, S. (2003). Reflections on past behavior: A self-report index of habit strength. *Journal of Applied Social Psychology*, *33*(6), 1313–1330. https://doi.org/10.1111/j.1559-1816.2003.tb01951.x

Wen, C. K. F., Schneider, S., Stone, A. A., & Spruijt-Metz, D. (2017). Compliance with mobile ecological momentary assessment protocols in children and adolescents: A systematic review and meta-analysis. *Journal of Medical Internet Research*, *19*(4), Article e132. https://doi.org/10.2196/jmir.6641

Wilson, T. D., Laser, P. S., & Stone, J. I. (1982). Judging the predictors of one's own mood: Accuracy and the use of shared theories. *Journal of Experimental Social Psychology*, *18*(6), 537–556. https://doi.org/10.1016/0022-1031(82)90072-5

Wood, W., Mazar, A., & Neal, D. T. (2022). Habits and goals in human behavior: Separate but interacting systems. *Perspectives on Psychological Science*, *17*(2), 590–605. https://doi.org/10.1177/1745691621994226

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: Thought, emotion, and action. *Journal of Personality and Social Psychology*, *83*, 1281–1297. https://doi.org/10.1037/0022-3514.83.6.1281

Zhang, Z., & Yuan, K.-H. (2018). *Practical statistical power analysis using WebPower and R*. ISDSA Press.