Determining the effects of training duration on the behavioral expression of habitual control in humans: a multilaboratory investigation

Eva R. Pool,[1,2,3,14] Rani Gera,[4,5,6,14] Aniek Fransen,[3,14] Omar D. Perez,[3,7,14]
Anna Cremer,[8,14] Mladena Aleksic,[3,9,14] Sandy Tanwisuth,[3,14] Stephanie Quail,[10,14]
Ahmet O. Ceceli,[11] Dylan A. Manfredi,[12] Gideon Nave,[12] Elizabeth Tricomi,[11]
Bernard Balleine,[10] Tom Schonberg,[4,5] Lars Schwabe,[8] and John P. O'Doherty[3,13]

[1]Swiss Center for Affective Sciences, University of Geneva, Geneva 1202, Switzerland; [2]Department of Psychology, University of Geneva, Geneva 1202, Switzerland; [3]Division of Humanities and Social Sciences, California Institute of Technology, Pasadena, California 91125, USA; [4]Sagol School of Neuroscience, Tel Aviv University, Tel Aviv 69978, Israel; [5]Department of Neurobiology, Tel Aviv University, Tel Aviv 69978, Israel; [6]School of Psychological Sciences, Tel Aviv University, Tel Aviv 69978, Israel; [7]Nuffield College Centre for Experimental Social Sciences Santiago, Faculty of Business and Economics, University of Santiago, Santiago 9170022, Chile; [8]Universität Hamburg, 20146 Hamburg, Germany; [9]Department of Social and Behavioral Sciences, University of Leiden, Leiden 2333, the Netherlands; [10]School of Psychology, University of New South Wales, Sydney, New South Wales 2052, Australia; [11]Department of Psychology, Rutgers University, Newark, New Jersey 07102, USA; [12]Marketing Department, the Wharton School of Business, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA; [13]Computation and Neural Systems Program, California Institute of Technology, Pasadena, California 91125, USA

It has been suggested that there are two distinct and parallel mechanisms for controlling instrumental behavior in mammals: goal-directed actions and habits. To gain an understanding of how these two systems interact to control behavior, it is essential to characterize the mechanisms by which the balance between these systems is influenced by experience. Studies in rodents have shown that the amount of training governs the relative expression of these two systems: Behavior is goal-directed following moderate training, but the more extensively an instrumental action is trained, the more it becomes habitual. It is less clear whether humans exhibit similar training effects on the expression of goal-directed and habitual behavior, as human studies have reported contradictory findings. To tackle these contradictory findings, we formed a consortium, where four laboratories undertook a preregistered experimental induction of habits by manipulating the amount of training. There was no statistical evidence for a main effect of the amount of training on the formation and expression of habits. However, exploratory analyses suggest a moderating effect of the affective component of stress on the impact of training over habit expression. Participants who were lower in affective stress appeared to be initially goal-directed, but became habitual with increased training, whereas participants who were high in affective stress were already habitual even after moderate training, thereby manifesting insensitivity to overtraining effects. Our findings highlight the importance of the role of moderating variables such as individual differences in stress and anxiety when studying the experimental induction of habits in humans.

[Supplemental material is available for this article.]

An accumulating literature suggests the existence of two distinct mechanisms for controlling instrumental behavior in mammals: a goal-directed mechanism, in which actions are selected with reference to the incentive value of an associated outcome, and a habitual mechanism in which action selection proceeds reflexively, underpinned by previously learned stimulus-response associations irrespective of current incentive value (Dickinson 1985; Balleine and O'Doherty 2010; Perez and Dickinson 2020). Elucidating the conditions under which goal-directed and habitual behavior arise has become a major research question, not only in the field of animal learning, but also in humans, as many aspects of human everyday experience can be profoundly influenced by the extent to which behavior is habitual or goal-directed (Ouellette and Wood 1998). Moreover, there has been increasing interest in the extent to which dysregulation in the balance of these two systems can contribute to aberrant behaviors in a number of psychiatric disorders, including addiction, obsessive compulsive disorder, eating disorders, and anxiety (Sjoerds et al. 2013; Alvares et al. 2014; Voon et al. 2015; Everitt and Robbins 2016; Gillan et al. 2016; Huys et al. 2016; but see Hogarth 2020 for a different account). Consequently, it is critical to gain an understanding of the

[14]These authors contributed equally to this work.
Corresponding authors: eva.pool@unige.ch, jdoherty@caltech.edu
Article is online at http://www.learnmem.org/cgi/doi/10.1101/lm.053413.121.

environmental and dispositional factors that can lead to the emergence of habitual and goal-directed behavior in humans.

The canonical assay for distinguishing goal-directed from habitual behavior in the laboratory is the outcome devaluation test (Adams and Dickinson 1981). In this procedure, an animal learns to form associations between instrumental actions (e.g., pressing on a lever) and particular rewarding outcomes (e.g., food pellets). Then, the rewarding outcome is devalued by feeding the animal to satiety (Dickinson et al. 1995; Balleine and Dickinson 1998) or pairing the outcome with gastric illness (Adams and Dickinson 1981). The key test of whether behavior is goal-directed or habitual arises when the animal is placed back in a test situation where the action it previously responded on is available but the delivery of the outcome is suspended (i.e., by testing under extinction). If the animal immediately decreases its instrumental action previously associated with the now devalued outcome, this indicates that behavior is goal-directed, in that it must be controlled by a representation of the response-outcome association. If, on the contrary, the animal persists in the instrumental action (relative to a control action whose outcome is not devalued), then behavior is argued to be habitual, or controlled by the stimulus–response association.

A seminal finding in rodents is that the amount of training can influence the behavioral manifestation of habits. Animals that are subjected to moderate training of an action remain predominantly goal-directed, manifested by a reduction in response rate to the action that has been associated with the devalued outcome. However, animals that are extensively trained (i.e., overtrained) become predominantly habitual, in that they fail to reduce responding on this devalued action (Dickinson et al. 1995; Holland 2004). Further experiments demonstrated that the response-outcome reward schedule that was in effect could also modulate the effect of training amount on devaluation sensitivity, showing that interval schedules tend to induce habits faster than ratio schedules for comparable training amount (Dickinson et al. 1983; Hilário et al. 2007; Wiltgen et al. 2012; Gremel and Costa 2013).

The animal findings spurred interest in determining whether similar differential effects of training occur in humans. Tricomi et al. (2009) adapted the instrumental free-operant procedure from the animal literature for use in humans, and deployed this procedure while participants were scanned with functional magnetic resonance imaging (fMRI). Using a specific satiety procedure similar to that used in rodents, these authors observed an effect of training on sensitivity to outcome devaluation, such that a group of participants exposed to longer training under an interval schedule were significantly less sensitive to outcome devaluation than participants exposed to only minimal training. This was interpreted as demonstrating that in humans, just like in rodents, extended training on an instrumental free-operant conditioning procedure can render behavior predominantly habitual.

These findings represented an important step toward the translation of animal findings to a human population. However, more recently, De Wit et al. (2018) reported the results of five experiments in humans where training manipulations did not produce any statistically significant effects on devaluation sensitivity. One of these experiments involved an appetitive instrumental task with an abstract cognitive devaluation without tangible outcomes, and therefore is not directly comparable with Tricomi et al. (2009). The investigators reported two additional studies involving an avoidance task where participants performed an instrumental response in order to avoid an unpleasant outcome. Nonetheless, avoidance procedures may be fundamentally different from appetitive procedures in terms of development and interaction between goal-directed and habitual processes. Therefore, such procedures are not straightforward to interpret, given that no studies to date have been able to clearly distinguish goal-directed and habitual

avoidance responding in free-operant training, which is the type of procedure used by Tricomi et al. (2009; see Fernando et al. 2014; Perez and Dickinson 2020 for possible interpretations of multiple systems involved in free-operant training). However, two of the five experiments run by De Wit et al. (2018) were described as a replication (outside the scanner) of the original Tricomi et al. (2009) where the authors also reported a failure to find evidence of an effect of the training amount on devaluation sensitivity. As is almost invariably the case with any replication, there were subtle and not so subtle differences between the replication paradigms and the original paradigm, rendering it challenging to ascertain whether the discrepancies between the results are due to differences between the paradigms, a false positive on behalf of the original Tricomi et al. (2009), false negatives on behalf of the two replication attempts (de Wit et al. 2018), or instead reflect the influence of other moderating variables (Camerer et al. 2018; Nave et al. 2020).

A key variable known to moderate the expression of habit and intuitive thinking is the level of stress (Dias-Ferreira et al. 2009; Schwabe and Wolf 2009; Soares et al. 2012; Starcke and Brand 2012; Otto et al. 2013; Margittai et al. 2016; Quaedflieg et al. 2019). One of the coauthors of this manuscript and his collaborators reported several findings showing that participants exposed to an experimental induction of stress exhibited an increase in devaluation-insensitive behavior indicative of stronger habitization than participants not exposed to the stress manipulation (Schwabe and Wolf 2009). Typically, these studies experimentally induce transient stress reactions such as by asking participants to put their arm in a bucket of icy water (Schwabe and Wolf 2009; Goldfarb 2019). However, the effects of individual differences in stress on habits are less well understood, particularly the effects on habitual behavior when a stressor is chronic (Arnsten 2015). Stress is conceived as a process of perceiving, responding, and adapting to threatening or challenging events (Lupien et al. 2007). Affective science distinguishes the process of stress elicitation and stress response (Lazarus and Folkman 1984). Stress is elicited by an event appraised as threatening to one's physiological and psychological integrity and exceeding one's available resources to successfully cope with it. The stress response includes a cognitive (e.g., worries), an affective (e.g., the feeling of negative affects), and a physiological component (e.g., activation of the hypothalamic–pituitary–adrenal axis) (Pool and Sander 2019). When stress is chronic and the stressors are repeated frequently over an extended time, the affective component of the stress response can mutate from a short affective episode triggered by a specific event, to a more general and diffuse mood, such as anxiety (Scherer 2005). Given the profound impact of chronic stress on brain and behavior, recently it has been proposed that the level of stress of participants should be taken into consideration in the design of human neuroscience studies (Goldfarb 2020). The moderating effect of stress could also be of particular interest in the context of the present study, where we aim to replicate a study originally performed inside an fMRI scanner in a behavioral testing room. Related to this, a recent large-scale investigation (Charpentier et al. 2020) demonstrated that participants enrolling in fMRI studies are lower in anxiety than participants enrolling in behavioral only studies perhaps because higher anxiety participants tend to avoid taking part in fMRI studies. Anxiety (and concomitant stress reactions) are thus potential candidates for accounting for the differential manifestation of habitual behavior in studies conducted inside versus outside the scanner.

Another important question that has not yet been explored with respect to the effects of stress on habits is whether stress acts to modulate the degree to which behavior transitions from goal-directed to habitual control as a function of training duration. A natural hypothesis in this regard is that stress could accelerate the

process of habit acquisition, such that the behavior of participants experiencing higher levels of stress shifts from goal-directed to habitual control more rapidly and therefore after a less amount of training.

To address the discrepancies between the existing studies on the effects of overtraining on habit formation and to examine the role of individual differences in stress and anxiety on the process of habit formation, we formed an international consortium on human habits (ICHB). Four laboratories undertook to run a preregistered replication of the original Tricomi et al. (2009) paradigm in >300 participants in total, outside the MRI scanner (see the Materials and Methods for the preregistration details). Each laboratory manipulated the amount of training (i.e., moderate or extensive) participants received for learning two instrumental actions leading to two different outcomes (i.e., sweet and salty snack) (see Fig. 1). After training, one of the two outcomes was devalued by feeding the participants to satiety and adaptation of the instrumental actions to the new values of the outcomes was tested under extinction. A subset of participants completed questionnaires measuring different facets of stress (Petrowski et al. 2012), anxiety (Spielberger et al. 1983), and impulsivity (Patton et al. 1995). We first report the results of our strict preregistered analyses from each one of the sites and then compare the size and the variability of the effect we found to the effects found in the Tricomi et al. (2009) and the de Wit et al. (2018) studies through a meta-analytical procedure. Finally, we take advantage of the large amount of data obtained from the preregistered protocol to further investigate the distribution of the effect of interest and explore the potential moderating effects of stress and anxiety on habit acquisition induced by extended training. This allows us to shed light on discrepancies between the previous findings by determining whether contradictory findings from the previous two studies can potentially be accounted for at least in part, by the influence of the moderating effect of stress on the process of habit formation.

Results

Results from the preregistered analysis

Manipulation check

The devaluation procedure significantly decreased the hunger level in each site (see Table 1 for detailed statistics; Fig. 2). We calculated the difference in the liking ratings of the two snacks used as outcomes (valued – devalued) and used this index as a dependent variable in a repeated-measures ANOVA. This analysis revealed a significant effect of Phase (predevaluation or postdevaluation) in each site (see Table 1 for detailed statistics) demonstrating that the decrease in pleasantness was significantly larger for the devalued food outcome compared with the valued food outcome (see Fig. 2). This shows that the selective satiation procedure for outcome devaluation was effective across all sites.

Outcome devaluation induced changes in each site

We hypothesized that the effect of the outcome devaluation procedure on instrumental response rates at test would be greater in the moderate training group compared with the extensive training group, because the extended training group was expected to exhibit a greater tendency to respond habitually, and hence to manifest an increased tendency to perform the action associated with the devalued outcome. To test this hypothesis, we calculated the difference in the average response rate per second during the free-operant task pre- and postcue devaluation (see Materials and Methods for more details; akin to Tricomi et al. 2009). This differential measure was used as the dependent variable in a 2 (cue: valued or devalued) × 2 (training: moderate or extensive) repeated measures ANOVA. We found that the interaction test did not reach significance in any of the five studies (see Fig. 3; Table 1 for detailed statistics). We found a main effect of cue in each one of the five studies, suggesting evidence for goal-directed behavior (see Table 1 for detailed statistics).

Results from the exploratory analysis

Meta-analytical comparison between our and previous effects

For a descriptive comparison of the effects found in each site of our study with the only other two existing studies published in the literature using the same identical paradigm (i.e., Tricomi et al. 2009;
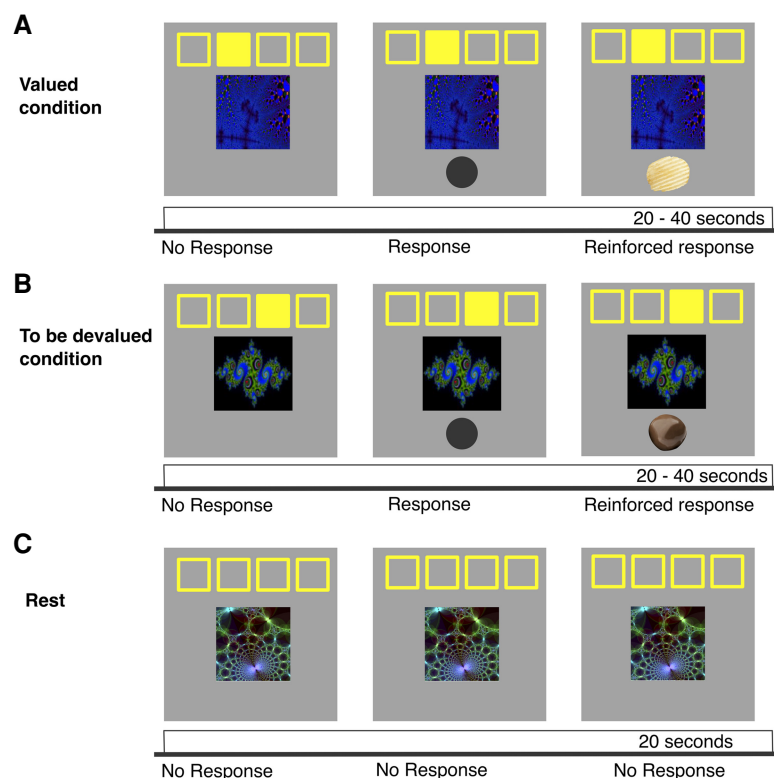
Figure 1. Illustration of the free-operant VI-10 task adapted from Tricomi et al. (2009). A fractal image appeared on the screen and stayed present throughout the block (20 or 40 sec), the filled-in yellow square indicated which button to press; the responses were self-paced. Each response activated a gray circle that stayed on the screen for 50 msec, every 10 sec on average a reward became available and the following response was reinforced either with a salty (A) or a sweet (B) snack. There were two cue–action–outcome combinations presented to a participant: one that remained valued throughout the experiment (A) and one that was devalued after training and before the extinction (B). (C) There was also a rest condition in which a third fractal was not associated with an action or a reward. The possible cue–action–outcome combinations were counterbalanced across participants.

Table 1. Detailed statistics of the manipulation check and the behavioral changes induced by outcome devaluation as function of the amount of training

| Site	F	P	η_p^2 [90% CI]
Main effect of *phase* on hunger ratings			
Pasadena 1	$F_{(60)} = 197.80$	<0.001	0.77 [0.68–0.82]
Pasadena 2	$F_{(57)} = 92.40$	<0.001	0.62 [0.49–0.71]
Hamburg	$F_{(62)} = 230.55$	<0.001	0.79 [0.71–0.84]
Sydney	$F_{(58)} = 309.26$	<0.001	0.84 [0.78–0.88]
Tel-Aviv	$F_{(58)} = 236.92$	<0.001	0.80 [0.73–0.85]
Main effect of *phase* on the snack liking index			
Pasadena 1	$F_{(60)} = 173.84$	<0.001	0.74 [0.65–0.80]
Pasadena 2	$F_{(57)} = 51.07$	<0.001	0.47 [0.32–0.59]
Hamburg	$F_{(63)} = 92.38$	<0.001	0.59 [0.47–0.69]
Sydney	$F_{(58)} = 87.72$	<0.001	0.60 [0.47–0.70]
Tel-Aviv	$F_{(58)} = 44.46$	<0.001	0.43 [0.27–0.56]
Cue × training interaction effect on the behavioral change [post − pre devaluation]			
Pasadena 1	$F_{(60)} = 0.73$	0.40	0.01 [0.00–0.09]
Pasadena 2	$F_{(57)} = 1.41$	0.24	0.02 [0.00–0.12]
Hamburg	$F_{(63)} = 0.91$	0.34	0.01 [0.00–0.10]
Sydney	$F_{(58)} = 0.19$	0.66	0.003 [0.00–0.07]
Tel-Aviv	$F_{(58)} = 0.42$	0.52	0.007 [0.00–0.08]
Main effect of *cue* on the behavioral change [post − pre devaluation]			
Pasadena 1	$F_{(60)} = 9.89$	0.003	0.14 [0.12–0.39]
Pasadena 2	$F_{(57)} = 14.87$	<0.001	0.21 [0.07–0.35]
Hamburg	$F_{(63)} = 17.97$	<0.001	0.22 [0.09–0.36]
Sydney	$F_{(58)} = 48.42$	<0.001	0.45 [0.30–0.58]
Tel-Aviv	$F_{(58)} = 5.39$	0.02	0.09 [0.01–0.22]

de Wit et al. 2018), we conducted a meta-analysis over these studies alongside our study illustrating the size and the variance of the effect of interest. For each one of the treatment groups in each one of the studies, we calculated an index of the behavioral adaptation to outcome devaluation by subtracting the behavioral change of the response rate per second [post − pre devaluation] in the valued condition from the behavioral change of the response rate per second [post − pre devaluation] in the devalued condition (referred to here as "behavioral adaptation index"). This behavioral adaptation index represents the difference of the behavioral change between the valued and the devalued cue induced by devaluation. A positive index value of the index is interpreted as goal-directed behavior, since the decrease in responding for the devalued cue would be larger compared with the valued cue. A value around zero is interpreted as habitual behavior, since the change induced by devaluation would be similar for the devalued and the valued cues. A negative index value represents an unexpected behavior, since the decrease would be larger for the valued cue compared with the devalued cue. We then calculated the effect size (standardized mean change [SMCC]) of the behavioral adaptation index ["cue valued post − cue valued pre" vs. "cue devalued post − cue devalued pre"] separately for the moderate training group

and for the extensive training group (see Fig. 4). The effect size of the behavioral adaptation index was larger for the valued than the devalued condition in both the moderate training group (SMCC = 0.54, 95% CI [0.38–0.69], $z = 5.26$, $P < 0.001$) and the extensive training group (SMCC = 0.41, 95% CI [0.27–0.56], $z = 5.64$, $P < 0.001$), suggesting evidence for goal-directed behavior in both groups. Even though descriptively the effect size of the behavioral adaptation index was slightly larger in the moderate training group compared with the extensive training group, this difference did not reach statistical significance ($Q_M = 1.41$, $df = 1$, $P = 0.23$). This was congruent with what we observed in the analysis by site and suggests that even across several studies the present paradigm does not reveal a statistically significant effect of training duration on the behavioral adaptation to devaluation.

Distributions of outcome devaluation induced change

The previous meta-analysis suggested that both groups adapt their behavior to devaluation, showing evidence for goal-directed behavior, independently of the amount of training (see Fig. 4). To further investigate this effect, we explored the distributions of the behavioral adaptation index ("cue valued post − cue valued pre" vs. "cue devalued post − cue devalued pre") representing the effect of interest, by pulling together the data obtained by each site of our study.

Visual inspection using density plots suggests that the distributions are likely multimodal, reflecting latent groupings in both conditions of the training manipulation (moderate or extensive) (see Fig. 5). We ran a finite mixture model on this data, and the results suggest that the distribution of the behavioral adaptation index is best explained by two latent clusters of participants (see
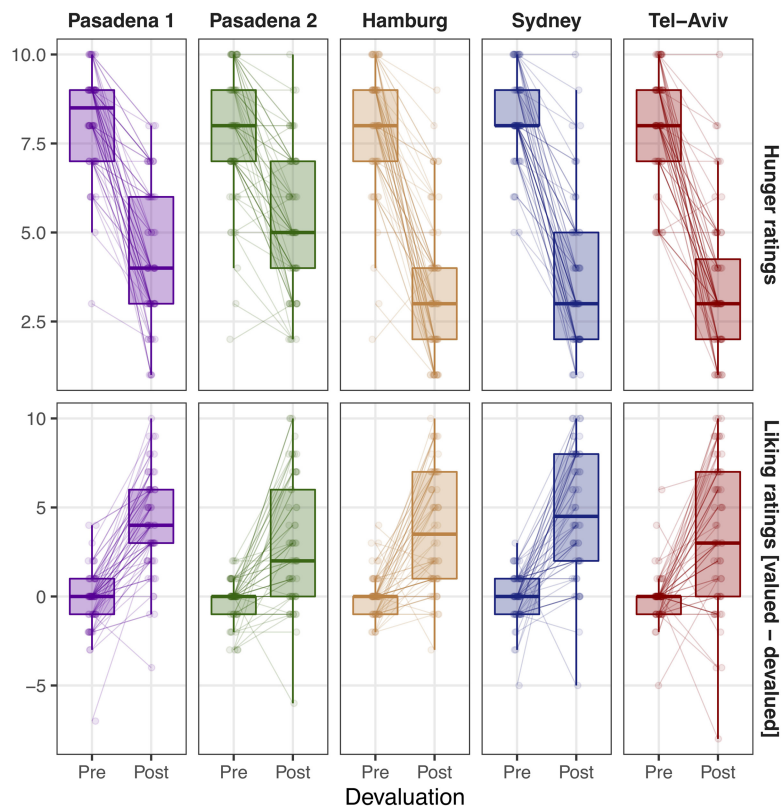
Figure 2. Manipulation check of the effectiveness of the outcome devaluation procedure. Ratings of hunger and the differential ratings of the snack liking [valued − devalued] before and after the outcome devaluation procedure in each one of the sites.
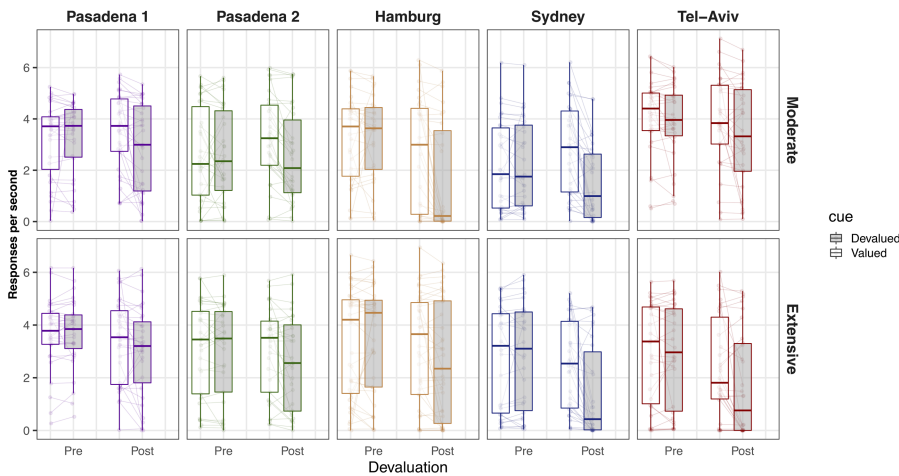
Figure 3. Average response rate per second during the valued and devalued cue before and after the outcome devaluation procedure in each one of the sites.

Table 2). The first cluster includes participants that adapted little or did not adapt their behavior to the outcome devaluation procedure ($n = 212$; outcome-insensitive). The second cluster included participants that changed their behavior after the outcome devaluation procedure ($n = 94$; outcome-sensitive). Consistent with the previous analyses, the frequency of outcome-sensitive and outcome-insensitive was descriptively similar in both the extensive training group (111 outcome-insensitive and 46 outcome-sensitive) and the moderate training group (101 outcome-insensitive and 48 outcome-sensitive). It is notable that participants who did not adapt their behavior to outcome devaluation constitute the majority of the sample in both groups (i.e., almost two-thirds of the participants in both groups). This is compatible with the hypothesis that the adaptation effect compatible with goal-directed behavior is driven by a smaller cluster of participants, while the lion's share of participants is in fact showing evidence of habitization on the task, in both the moderate and extensive training groups.

Analysis of the moderating effects of individual differences relating to stress, anxiety, and impulsivity on habit formation as a function of training duration

In a follow-up analysis we tested for the presence of potential moderating variables on our effects of interest. A subset of participants ($n = 199$) fully completed three questionnaires measuring stress (nine subscales), impulsivity (three subscales), and anxiety (a single composite score) (see the Materials and Methods for details). Some participants did not complete these questionnaires because they were not administered systematically at each site, and were administered to only half of the participants in an ex-

periment run at one of the other sites (see the Materials and Methods for details). We ran an analysis to test whether these variables could moderate the effect of the amount of training on the acquisition and expression of habits. Given that these subscales were highly correlated, we conducted a factorial analysis on the subscales of the questionnaires to extract factors that could be later entered as predictors in the statistical model testing the effect of training on devaluation sensitivity (see Gillan et al. 2016; Patzelt et al. 2019 for a similar approach). The analysis suggested a four-factor solution that we labeled "Stress Work," "Impulsivity," "Stress Social," and "Stress Affect" (see Table 3; see the Materials and Methods for details).

We then entered these factors into a multilevel model where we tested the effect of training on the sensitivity to devaluation through the interaction between: Cue (valued or devalued), Phase (pre- or postdevaluation), and Training (moderate or extensive). The analysis revealed a significant interaction between Cue, Phase, Training, and the "Stress Affect" factor ($\beta = -0.26$, SE = 0.09, 95% CI [-0.46, -0.07], $P = 0.007$). Simple slopes follow-up tests revealed that the interaction between cue, value, and group was positive and significant in participants with lower (-1 SD)
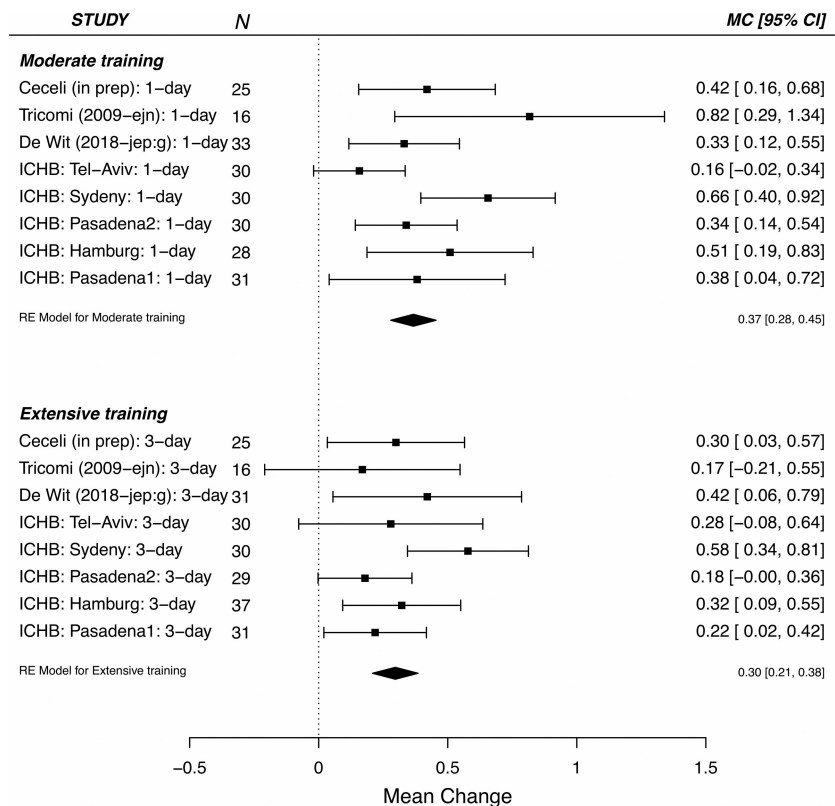
Figure 4. Forest plot illustrating the effect size (standardized mean change [SMCC]) and the 95% confidence interval (95% CI) of the behavioral adaptation index ["cue valued post – cue valued pre" vs. "cue devalued post – cue devalued pre"] in the moderate training group (1 d of training) and the extensive training group (3 d of training) from several studies using the same paradigm ($N = 402$).

Table 2. Model fit statistics for finite mixture model

k	logLik	AIC	BIC	Cluster sizes
1	−328.01	662.02	673.19	306
2	−243.60	501.20	527.27	212, 94
3	−240.34	502.79	543.75	218, 72, 16
4	−239.85	509.71	565.56	61, 26, 198, 21
5	−239.26	516.51	587.26	7, 22, 30, 196, 51

(k) Number of clusters, (logLik) log likelihood, (BIC) Bayesian Information Criterion, (AIC) Akaike Information Criterion.

levels of "Stress Affect" ($\beta = 0.38$, SE = 0.13, 95% CI [0.10, 0.64], $P = 0.006$), whereas it was not significant (with a negative point estimate) in participants with a higher (+1 SD) level of "Stress Affect" ($\beta = -0.15$, SE = 0.14, 95% CI [−0.41, 0.12], $P = 0.28$) (see Fig. 6). We did not find evidence for a significant interaction between factor "Stress Work" ($\beta = -0.14$, SE = 0.10, 95% CI [−0.33, 0.05], $P = 0.15$), "Stress Social" ($\beta = -0.17$, SE = 0.10, 95% CI [−0.36, 0.02], $P = 0.08$), and "Impulsivity" ($\beta = -0.03$, SE = 0.10, 95% CI [−0.22, 0.17], $P = 0.77$) and the effect of interest (i.e., the interaction between Cue, Phase, and Training).

Robustness checks on the analysis of the moderating effects of individual differences

To ensure the robustness of our factor analytical conclusions, we ran the factor analysis using two different methods. The first of these approaches is the one reported here (see the Materials and Methods), and the second is reported in the Supplemental Material (see strategy 2, Supplemental Table S1; Supplemental Fig. S1). Both methods yielded similar factor structures, and the statistical analyses support the same conclusions about the role of affective stress on moderating habit formation. For yet another robustness check, we ran a similar analysis but this time forgoing the factor analysis approach and instead using the relevant specific subscales for anxiety and chronic worrying (see strategy 3, Supplemental Table S2; Supplemental Fig. S2). Once again, the same results were found, thereby supporting the overall robustness of our factor analytic conclusions.

Discussion

The main objectives of this preregistered multilaboratory investigation were first to determine the extent to which training could produce an increase in habitual responding, and second, shed some light on contradictory findings from previous studies investigating this critical question by investigating the role of potential moderating variables such as stress, anxiety, and impulsivity. Our findings suggest that the process of habit acquisition appears to be modulated by individual differences in the level of affective stress. An effect of overtraining on the sensitivity of outcome devaluation was observed in participants reporting low levels of affective stress but it was not present in participants reporting a higher level of affective stress.

In the preregistered part, we used the paradigm initially used by Tricomi et al. (2009) to distinguish habits from goal-directed action based on the sensitivity of the instrumental action to outcome devaluation (Adams and Dickinson 1981). We did not find statistical evidence supporting an effect of the amount of training on the development of habitual behavior. Instead, when looking at the mean differences between groups (moderate and extensive training) and devaluation (valued and devalued trials), our findings appear to be congruent with the findings of de Wit et al. (2018), showing a main effect of devaluation, but no significant interaction with the amount of training. This supports the interpretation of the presence of goal-directed behavior independently of the amount of training. However, a closer inspection of the distributions of the behavioral adaptation to outcome devaluation revealed that the mean differences between the valued and devalued conditions were driven by a smaller proportion of participants. The majority of the participants did not adapt their behavior flexibly to the devaluation procedure in both the moderate and the extensive training group. This occurred despite the fact that the outcome devaluation procedure was highly effective in both groups in terms of reducing the pleasantness ratings for the foods. This might indicate that, rather than there being a prevalence of strongly goal-directed behavior in our sample, there was in fact a strong prevalence of habitual behavior in both the moderate and extensive training groups, even if a minority of participants nevertheless exhibited sufficient devaluation sensitivity to yield a significant effect of devaluation on instrumental actions overall. Note that also de Wit et al. (2018) reported responding for the devalued outcome after devaluation in each one of their experiments, which is congruent with the presence of habitual behavior in a subset of participants.

Interestingly, we found a significant moderating effect of a factor "stress affect" on the production of habits as a function of training duration. This critical factor reflects the affective components of chronic stress such as anxiety, worries, isolation, and discontent. Specifically, those participants high in affective stress appeared to manifest outcome-insensitive behavior even after moderate training. More precisely, those participants who were low in affective stress appeared to be devaluation sensitive after shorter training, while after longer training they appeared to transition to habitual behavior. While we emphasize the exploratory nature of these findings, we note that they resonate with an existing literature demonstrating a key moderating role for anxiety and stress on the behavioral expression of habits in both humans and rodents (Packard 1999; Schwabe et al. 2008, 2011; Dias-Ferreira et al. 2009; Schwabe and Wolf 2009, 2010; Soares et al. 2012; Otto et al. 2013; Goldfarb et al. 2017; Quaedflieg et al. 2019; Hartogsveld et al. 2020). Congruently with our findings, recent empirical evidence suggests that stress could accelerate the shift from goal-directed behavior to habit performance even after

Table 3. Loading onto factor 1 "stress work," factor 2 "impulsivity," factor 3 "stress social," and factor 4 "stress affect"

	Stress work	Impulsivity	Stress social	Stress affect
Anxiety composite score	0.139	0.082	−0.005	**0.626**
BIS attentional	0.120	0.419	−0.006	0.268
BIS motor	−0.064	**0.604**	0.108	−0.004
BIS nonplanning	0.013	**0.857**	−0.018	−0.024
TICS chronic worrying	0.359	−0.046	0.113	0.436
TICS excessive demands at work	**0.941**	0.064	−0.040	0.058
TICS lack of social recognition	0.184	−0.004	0.504	0.190
TICS pressure to perform	0.132	−0.123	**0.667**	0.182
TICS social isolation	−0.086	−0.036	0.025	**0.866**
TICS social overload	−0.062	0.045	**0.911**	−0.091
TICS social tensions	0.112	0.139	0.513	0.195
TICS work discontent	0.171	0.176	0.111	0.507
TICS work overload	**0.743**	−0.099	0.151	−0.124

The top two scores for each factor are highlighted in bold.
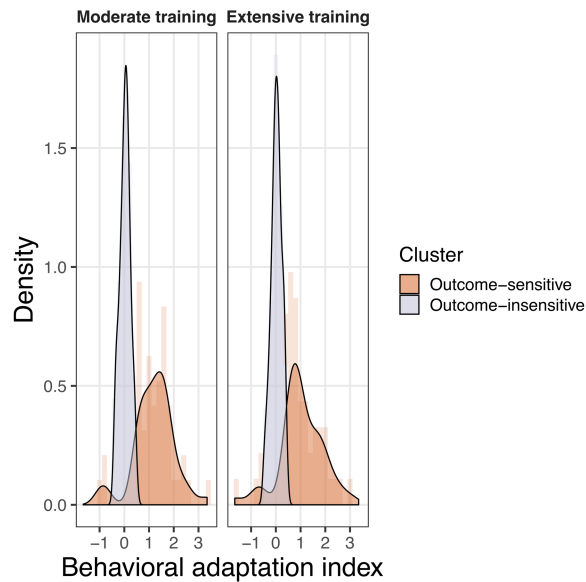
Figure 5. Distributions of the behavioral adaptation index to outcome devaluation in our study (*n* = 306). The behavioral adaptation index ["cue valued post – cue valued pre" vs. "cue devalued post – cue devalued pre"] calculated on the response rate per second during the free operant task is displayed in the moderate training and extensive training groups. Two clusters of participants were identified: Outcome-sensitive (*n* = 94), which modified their behavior after devaluation compared with before devaluation, and Outcome-insensitive (*n* = 212), which did not modify their behavior after devaluation as compared with before devaluation.

moderate training (Meier et al. 2021). However, distinct from most studies investigating the impact of stress on habit formation, we sorted people based on pre-existing individual differences in self-reported chronic stress and anxiety as opposed to using an experimental stress induction procedure to test the effect of overtraining in interaction with these individual differences. The few studies that sorted participants based on pre-existing individual differences are also congruent with our results. They showed that individual differences in chronic stress (Schwabe et al. 2008) or in lifetime stress history (Goldfarb et al. 2017) were associated with use of habit-like stimulus–response strategies.

Previous empirical work has investigated the influence of self-reported anxiety on the propensity to rely on habits rather than goal-directed actions in moderately trained participants, using a variety of paradigms (free-operant conditioning, the two-step task, and the slip-of action task). Findings are mixed: Two studies found no influence of anxiety on the impairment of goal-directed strategies (Gillan et al. 2016, 2021), while other investigations found a relationship between anxiety and the propensity to rely more on habits (Snorrason et al. 2016; Ersche et al. 2017) or less on goal-directed strategies (Patzelt et al. 2019). A single study using a free-operant design like the one used by Tricomi et al. (2009) also found a direct relationship between anxiety and habitual behavior (Alvares et al. 2014). Please note that our current findings about the effects of stress on moderating overtraining effects on habits are purely correlational, in that we have no evidence about the causal role of stress on the acquisition or expression of a habit.

Importantly, our findings on the moderating effect of affective stress on habit formation with extended training could shed some light into the discrepancies between the results reported by Tricomi et al. (2009) and de Wit et al. (2018). Perhaps the most obvious difference between the original study and the replications is

that the original Tricomi et al. (2009) study was conducted inside an fMRI scanner, whereas the de Wit et al. (2018) replications and our replications were conducted in a behavioral testing suite. A recent large-scale investigation (Charpentier et al. 2020) demonstrated that participants enrolling in fMRI studies are lower in anxiety on average than participants enrolling in behavior-only studies. This points at a potential selection bias: Higher-anxiety individuals are likely more reluctant to volunteer for experiments that can be potentially stressful, such as fMRI studies. It is thus possible that participants in the Tricomi et al. (2009) fMRI experiment could have been less anxious individuals compared with the participants in the behavioral replications. In our findings, participants reporting lower levels of affective stress did show an effect of training duration on habit formation (consistent with Tricomi et al. 2009), whereas participants reporting higher levels of affective stress did not (consistent with de Wit et al. 2018). This bias could explain why the original study had found an overtraining effect on devaluation sensitivity inside the scanner, whereas subsequent studies conducted outside the scanner did not.

Irrespective of the possible differences between fMRI and behavioral studies, the finding of a potential moderating effect of the affective component of chronic stress on habit formation is theoretically and clinically relevant, implying that longer training may indeed elicit habits in humans, yet the amount of training necessary to induce habits varies according to the individual level of affective stress. It is notable that among the different factors related to chronic stress, such as social factors or factors related to work overload, the determining factor in our sample was related to the affective factors of stress such as anxiety and worries. The Trier Inventory for Chronic Stress (TICS) (Petrowski et al. 2012), conceptualizes stress as an interaction between the environment (e.g., with high demands or lack of wanted positive events) and the individual (e.g., resources to cope with the event). It seems plausible that the "stress affect" factor derived from our factor analysis is (relative to the other factors) most sensitive to the gap between the stressful event and the appraisal of the individual's resources as being insufficient to successfully cope with such an event. This could have important implications for testing the impact of training extension on habit formation. So far, empirical work investigating habits in humans usually measures the balance between the goal-directed and habitual system in a specific individual by fitting their behavior to model-based or model-free algorithms rather than tracking the shift from goal-directed to habitual control with extended training (Radenbach et al. 2015; Voon et al. 2015; Gillan et al. 2016; Patzelt et al. 2019). As a result, findings on the experimental induction of habits by extended training in humans remain limited and contradictory (Tricomi et al. 2009; de Wit et al. 2018; Luque et al. 2019; but see Hardwick et al. 2019 for new promising evidence).

We did not find any significant moderating effect of impulsivity on habit formation in our paradigm, which differs from previous reports (Gillan et al. 2016) finding that impulsivity was associated with reduced model-based control. Impulsivity has been shown to be a multidimensional construct that reflects a combination of separable psychological dimensions such as the tendency to experience strong reactions (urgency), lack of premeditation, or sensation seeking (see Whiteside and Lynam 2001). Perhaps tools assessing the multifaceted nature of impulsivity could lead to more congruent results.

It is important to note that there are a number of limitations to this study. First, there was a discrepancy across sites in the version of the anxiety questionnaires used (two sites used state and one site used trait). While these measures are likely strongly correlated, they reflect different underlying constructs: It will be important to obtain clarity on which form of anxiety loads more strongly onto the affective stress factor influencing habit formation.
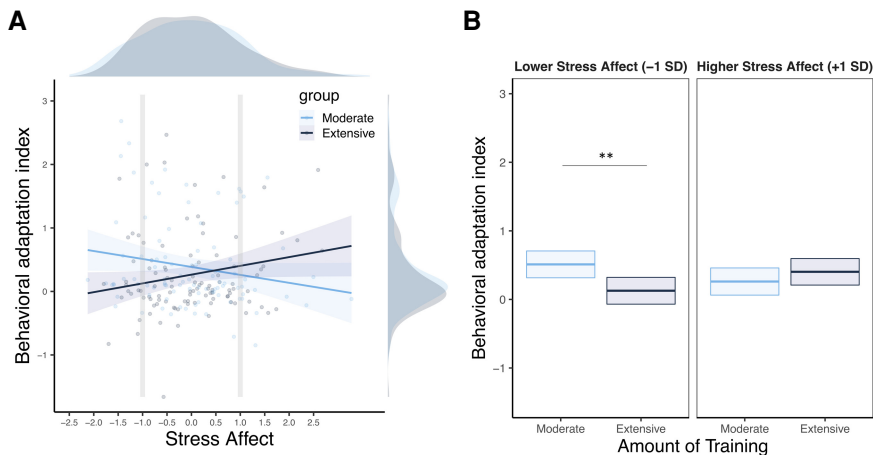
Figure 6. (*A*) Behavioral adaptation index ([“cue valued pre − cue valued post” vs. “cue devalued pre − cue devalued post”] calculated on the response rate per second during the free operant task, *n* = 199) as a function of the level on the “Stress Affect” factor in participants that received either a moderate or an extensive amount of training. Shaded areas indicate the 95% CI. (*B*) Mean adjusted behavioral adaptation index to moderate versus extensive training as a function of lower (−1 SD) and higher (+1 SD) level of the “Stress Affect” factor.

Second, there were other variations in experimental procedures across studies (see the Materials and Methods for details). Although we do not think these made a substantive difference to our findings, it would be useful to ensure these procedures are as identical as possible across sites in future studies of this kind in order to minimize extraneous sources of variance across sites. Third, the findings on the moderating effect of affective stress were from an analysis that could only be conducted on a subsample of our participants because the questionnaire data was obtained consistently only in a subset of the participants participating in this study. This emerged in part because of the complexity in organizing a large-scale multisite study of this kind in which different laboratories varied in their procedures, and because we elected to include questionnaire measures of individual differences only at a relatively advanced stage in the project. As we specifically, and the field more generally, gain greater experience in coordinating this kind of large-scale collaboration across laboratories, in the future we expect that improved coordination and communication will lead to greater consistency in the methods and measures used.

An important insight arising from our study concerns the sensitivity of the specific experimental paradigm we used to measure training effects on habits in humans that was first deployed by Tricomi et al. (2009). In the present study we found devaluation-insensitive behavior in the majority of participants irrespective of whether they were in the moderately trained or overtrained groups, which is consistent with the possibility that the majority of our participants were actually habitual after even moderate training. This suggests that in actuality, the Tricomi habit paradigm might be too prone to induce habits, doing so rapidly in the majority of participants after only moderate training. Thus, the experimental challenge in obtaining evidence for the effect of training duration on habit formation in humans is one of a need to deploy a paradigm that produces stronger goal-directed behavior for limited training as opposed to one that is more prone to develop habitual control of behavior independently of training extension. If the majority of participants in both groups are rapidly habitized after moderate training, then a longer training extension will not be able to affect devaluation sensitivity, as there is no experimental variance in behavior left to modulate it. This effect may be compounded by individual differences: It was only in those participants low in affective stress in the present study, that we saw an

overtraining effect because only those participants were sufficiently goal-directed at the beginning of training to show the change of behavioral control as training progressed.

Another potential explanation for why this paradigm is so sensitive to habits as opposed to goal-directed behavior is the employment of a variable interval schedule to determine the response-outcome contingency. Variable interval schedules are known to produce rapid habit formation in rodents (Dickinson et al. 1983; Wiltgen et al. 2012; Gremel and Costa 2013; Perez and Dickinson 2020). The present findings suggest that rapid habit formation following a variable interval schedule might occur in humans too; however, more evidence in further studies is needed for this interpretation by directly contrasting different schedules on the propensity to habit formation. In particular, ratio as opposed to interval schedules is known to more robustly produce goal-directed behavior in rodents (Dickinson et al. 1983). Therefore, an important future direction is to use a ratio schedule as opposed to an interval schedule in order to examine overtraining effects in humans.

Recently, it has been claimed that for interpreting behavior as habitual or goal-directed it is crucial to first show that (1) the devaluation has been effective, and (2) that the test administered after devaluation is sensitive enough to reflect goal-directed behavior, if that is present (De Houwer et al. 2018; Moors et al. 2017). If these two sensitivity criteria are not satisfied, the interpretation of behavior during the test phase is argued to be ambiguous. Although we found evidence that the devaluation procedure was strongly effective in the present study ruling out the possibility of the first of these concerns, we cannot completely exclude the possibility that the test phase in the present paradigm is adequately sensitive to goal-directed behavior. For example, participants might have been aware that even if they collect food outcomes that they no longer value by performing instrumental responses for those outcomes, they cannot be forced to consume those foods at the end of the experiment. Alternatively, the shift to extinction conditions may have reduced the transfer of information encoded in training to the test, something that may also have interacted with stress. Finally, obtaining those undesired foods through responding presents no tangible cost other than the cost of responding, which is rather miniscule for an action such as key pressing. Thus, they may have been indifferent as to whether to respond or not for a now devalued outcome. Therefore, another direction for improvement of the paradigm would be to ensure that decreased responding following devaluation is clearly consistent with participant’s interests, aligning incentives on the task with performance under the case where behavior is goal-directed in a manner that is clearly understood by the participants. One way to achieve this goal would be to impose an explicit cost of responding in the task. Such cost should also help participants encode the reward schedule in effect during training and perform at a rate that reflects the properties of the response-outcome contingency (Reed 2001; Perez and Dickinson 2020; Perez and Soto 2020).

In our task, we used sensitivity to outcome devaluation as an index of goal-directed or habitual behavior. This manipulation does not allow one to disentangle whether the expected transition from goal-directed behavior to habits is due to (1) an increase of habitual control, (2) a decrease of goal-directed control, or (3) both.

Even though theoretically these two mechanisms can be measured separately at the behavioral level they are deeply intertwined and very hard to distinguish. To date this remains one important challenge that needs to be effectively addressed (Balleine and Dezfouli 2019) and that is particularly important to the study of stress effects, which are known to impair goal-directed processes (Devilbiss et al. 2017). Indeed, recent views of the interaction of goal-directed and habitual action argue that the contribution of the action–outcome association to performance generally declines over the course of overtraining as habitual control increases (Perez and Dickinson 2020).

A final limitation is that the tests we performed on the moderating effects of individual differences in the present study are exploratory as opposed to confirmatory, unlike the test for the main effect of overtraining on habit formation that was fully preregistered. We hope that future studies would further attempt to replicate the present findings about the moderating effects of stress and anxiety on the effect of training duration using a fully preregistered approach focused on these moderating effects. A subsequent confirmatory study on this question would help advance confidence in these exploratory findings. Critically, future studies investigating these individual differences should aim at recruiting participants that are more diverse and representative of the general population than our sample, which included a very large proportion of students (Simons et al. 2017).

In conclusion, we could not find evidence for a main effect of the amount of training on habit formation tested by the sensitivity to devaluation in the present study. Instead, the large majority of the participants of our sample showed little sensitivity to outcome devaluation, expressing habit-like behavior already after even moderate training. However, our findings suggest that factors related to stress and anxiety can accelerate habit formation, thereby exerting a moderating effect on training duration in the expression of human habitual behavior.

Materials and Methods

Participants
A total of 327 participants were tested (see Table 4). Seventeen participants were excluded based on preregistered criteria (see https ://aspredicted.org/5ns9z.pdf), three participants were excluded because of technical problems (i.e., the door of the experimental room was not closed in one case and participants stopped doing the task in two cases), and one was excluded for having extreme values in the free-operant task (i.e., an action–outcome devaluation effect larger than five interquartile range from the median).

A total of 306 participants were included in the primary analysis pooled across all experiments, which tested for an overall effect of the training amount on devaluation sensitivity and a total of 199 participants were included in the secondary analysis, testing for individual difference effects moderating the effect of the training amount on devaluation sensitivity.

Please note that a technical problem in the data collection in Hamburg caused a subset of the participants to receive more food outcomes for consumption than would be expected given the reward contingencies. We ran the analyses with and without these participants and did not find a significant impact on the outcome of the analyses; we therefore included all participants in the final statistical analyses.

Materials

Procedure
The experimental procedure involved five main parts: (1) a snack selection phase, (2) a free-operant task phase, (3) an outcome devaluation procedure phase, (4) extinction test phase, and (5) a questionnaire phase. Participants were divided into two experi-

Table 4. Summary of the demographic information by site

Site	Age M (SD)	Female (N)	Male (N)	Nonbinary (N)
Pasadena 1	26.22 (8.41)	38	23	1
Pasadena 2	-	36	23	-
Hamburg	25.07 (4.30)	36	29	-
Sydney	23.15 (3.89)	42	18	-
Tel-Aviv	25.15 (4.26)	42	18	-

(M) Mean, (SD) standard deviation.

mental groups: a moderate and an extensive training group. The moderate training group underwent the experimental procedure in a single day, whereas the extensive training group underwent the experimental procedure in three consecutive days (see Fig. 7).

Snack selection phase
The main dependent variable consisted of forced choices between snacks. Participants were presented with a selection of individual pieces of six snacks divided in two categories: sweet and savory. They were asked to taste each sample and choose their favorite savory snack and their favorite sweet snack.

Free–operant training phase
The paradigm of the free-operant task was identical to the paradigm used in Tricomi et al. (2009) (see Fig. 1), in which participants' responses were rewarded. These responses were rewarded with two possible food outcomes (a sweet and a savory snack) to be consumed following the task. One group of participants performed two training sessions on 1 d, whereas a second group of participants performed four training sessions each day for 3 d. Each session was divided into 12 task blocks and eight rest blocks. During the task blocks, a fractal image (i.e., a cue) was shown on the screen, along with a schematic indicating which button to press, and stayed present throughout the block (20 or 40 sec). Participants were instructed to press the indicated button as often as they like; after each button press (i.e., a response) two possible outcomes could appear on the screen: either a gray circle (for 50 msec), indicating no reward, or a picture of a sweet snack or savory snack, indicating a food reward corresponding to the picture (for 1000 msec). Rewards were delivered with a variable interval 10 sec schedule (VI 10 sec): Each second there was a chance that a reward would be available upon a button press, so that on average a reward became available every 10 sec. Different fractals and response keys were paired with the two outcomes, and these cue–action–outcome associations remained consistent throughout the experiment. A third fractal indicated a rest block, during which participants were instructed not to respond. The order of the block was pseudorandomized, with no block type occurring twice in a row. Following the final session of training, one of the two food outcomes was devalued.

Outcome devaluation procedure
The devaluation procedure occurred through the selective satiation of one of the two food outcomes used in the free-operant task. As a control variable the amount of consumed food was weighed. As a manipulation check of the effectiveness of the devaluation procedure, ratings of hunger and pleasantness were collected prior to each day's training session and following the devaluation procedure. The main dependent variable was Likert-scale ratings of hunger (1, very full; 10, very hungry) and pleasantness (−5, very unpleasant; 5, very pleasant).

Extinction phase
To test for the effects of the devaluation procedure on participants' behavior in the absence of any further experience with the outcome, an extinction test was administered where the same responses were available but reward delivery was suspended. The
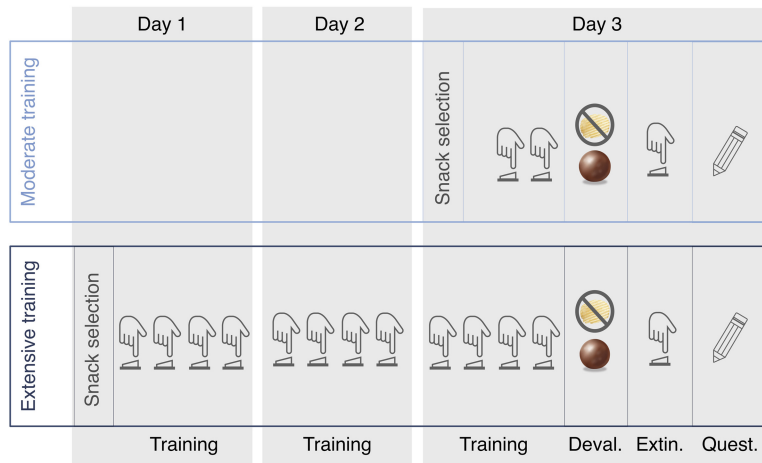
Figure 7. Illustration of the experimental procedure adapted from Tricomi et al. (2009). Participants were divided in two groups: moderate and extensive training group. Participants of the extensive training groups came 3 d to the laboratory where they received four learning sessions each day, whereas participants from the moderate training group came only 1 d to the laboratory where they received two learning sessions. After selecting their favorite sweet and salty snack (i.e., snack selection), participants received either a moderate or an extensive training where different actions were associated with their favorite sweet and salty snack, respectively (i.e., Training). Then one of the two snacks was devalued via feeding to satiety on that snack (i.e., Deval.) and participants underwent an extinction test administered under extinction (i.e., Extin.). Finally participants completed a series of questionnaires (i.e., Quest.).

extinction test was composed of six task blocks and three rest blocks (i.e., three blocks per condition). The extinction test was implemented in the same manner as for the free-operant training sessions: The fractal images and schematic indicating which button to press stayed present throughout the block (20 sec), and after each response a gray circle was displayed for 50 msec; however, no rewards were actually delivered. Each site strictly followed this protocol, the task scripts were the same for every site and the instructions were exactly the same translated into the relevant language. The stimulus presentation and behavioral data acquisition were implemented in Matlab (The Mathworks, Inc.) with the psychophysics toolbox extensions (Brainard 1997; Pelli 1997), questionnaires were administered either online or in paper and pencil depending on the site (the scripts for stimulus presentation and data acquisition, the instructions and the written protocol closely followed by each site are openly available). There were, however, two methodological variations. First, in four sites participants received the food outcomes to be consumed after each session rather than at the end of the training day (Hamburg, Pasadena2, Sydney, and Tel-Aviv), while in one other site (Pasadena1) the food outcomes to be consumed were presented at the end of the training session the same as Tricomi et al. (2009). Second, one of the sites (Pasadena2) added a number of modifications that aimed to match the conditions experienced inside an fMRI scanner. These included having participants lie in a supine position, playing fMRI sequence noise via a speaker positioned below the scanner bed through which typical MRI sequence noise was played at 60 db, and monitoring participants eye and head movements while requiring them to remain as still as possible.

Questionnaire phase

The questionnaires were administered at the end of the experimental procedure. The TICS (Petrowski et al. 2012) and the State/Trait Anxiety Inventory (the state version was run in Hamburg and Tel-Aviv; the trait version in Pasadena1 and Pasadena2; Spielberger et al. 1983) were used to assess stress and anxiety respectively. The Barratt Impulsivity Scale (BIS-11 for Pasadena1, Pasadena2, and Tel-Aviv; BIS-15 for Hamburg) (Patton et al. 1995). The different subscales of the BIS were standardized and used as indexes to reflect motor impulsivity, attentional impulsivity, and nonplanning. Moreover, some participants also completed

the Obsessive Compulsive Inventory (OCI-R; only in Pasadena1, Pasadena2, and Tel-Aviv) (Foa et al. 2002) and the Beck Depression Inventory (BDI; only in Hamburg) (Beck et al. 1988). As there was an inadequate sample size to draw useful conclusions about the effects of those measures given that these additional questionnaires (BDI and OCI-R) were not administered consistently across all groups who collected questionnaire data, we did not focus on them any further. However, the data are made available to the community for further interest.

We decided to collect questionnaire data only after the completion of data collection for the first study in Pasadena (Pasadena1), but before data collection for the other studies. For this reason, we recontacted the participants from the Pasadena1 study and obtained questionnaire data from those participants 1 mo after the original data was acquired. Only ~50% of participants from that study subsequently agreed to return the questionnaire data when contacted. Questionnaire data was not collected at the Sydney site. Given that the STAI, TICS, and BIS were the only questionnaire measures consistently collected across the three sites at which questionnaire data were collected, we focused our analyses on those specific questionnaires. It should be noted that the STAI measures also differed between sites: two sites collected the state subscale and one site collected the trait subscale. Because state and trait anxiety measures are highly correlated (Spielberger 2013), we reasoned that we would still be able to detect meaningful variance related to anxiety by pooling the data (standardized) across sites, even if it could not be unambiguously attributed to state or trait anxiety effects per se.

Statistical analyses

Statistical analyses were performed with RStudio (version 1.1.442; RStudio Team 2016). Statistical analyses were divided into two phases: preregistered and exploratory. The preregistered analyses are a strict replication of the analysis reported by Tricomi et al. (2009) and were performed on each site separately, whereas the exploratory analyses focused on a comparison between our data and the data existing in the literature (meta-analysis), an exploration of the distribution of the effect of interest by merging the data from all sites (Cluster analysis) and a test of variables moderating the effects of interest by merging the data from all sites (Factor analysis and multilevel analysis). In some of the analyses we defined a pre- and postdevaluation window to test for behavioral change. Akin to Tricomi et al. (2009) we used response rate per second during the last block of training as measure of predevaluation (i.e., six blocks per condition) and the extinction test as measure of postdevaluation (i.e., three blocks per condition). This allowed us to control for baseline differences in the response rate when performing the tests of interest during the extinction phase.

Preregistered analyses

ANOVAs

We used the afex (Singmann et al. 2015). Adjustments of degrees of freedom using Greenhouse-Geisser correction were applied when the sphericity assumption was not met. Partial eta squared (η_P^2) and their 90% CI are reported as estimates of effect sizes for the ANOVAs. We additionally computed the Bayes factor (BF$_{10}$) quantifying the likelihood of the data under the alternative hypothesis

relative to the likelihood of the data under the null hypothesis using Bayesian ANOVAs (e.g., Rouder et al. 2012).

Exploratory analyses

Meta–analysis

To descriptively compare the effects we obtain in each site to the two existing studies in the literature (Tricomi et al. 2009 and de Wit et al. 2018), we conducted a small meta-analysis using the metaphor package (Viechtbauer 2010). We included the data from Tricomi et al. (2009), the data from de Wit et al. (2018), and our multilaboratory data set. We did not do a systematic search of all possible studies using different kinds of overtraining procedures, since the main objective of the meta-analysis was to provide a descriptive comparison of the effect across the experimental studies using the exact same paradigm as the one we used in the present study. We calculated an index of the behavioral adaptation to outcome devaluation by subtracting the behavioral change [post – pre devaluation] in the valued condition from the behavioral change [post – pre devaluation] devalued condition (i.e., "behavioral adaptation index"). We then extracted the effect size of the behavioral adaptation index ["cue valued post – cue valued pre" vs. "cue devalued post – cue devalued pre"] and its variability (95% CI). Note that, throughout, "cue" refers to the "cue + response" pair that remained valued or that was devalued. The effect size we extracted was the standardized mean change using change score standardization (SMCC) (Morris and DeShon 2002). If after devaluation the behavioral response was larger for the valued than the devalued condition the effect size was given a positive sign. We compute the effect size in each condition (i.e., moderate and extensive training) using a random effect model (RE), whereas we compute the moderator analysis (i.e., effect of the amount of training on the size of the effect) using a fixed-effect model (Borenstein et al. 2011).

Cluster analysis

For the clustering analysis we applied the FlexMix clustering algorithm (Leisch 2004) on the behavioral adaptation index (cue valued post – cue valued pre vs. cue devalued post – cue devalued pre) as a function of the amount of training. The behavioral adaptation index represents the difference of the behavioral change between the valued and the devalued cue induced by devaluation. A positive index value of the index is interpreted as goal-directed behavior, since the decrease in responding for the devalued cue would be larger compared with the valued cue. A value around zero is interpreted as habitual behavior, since the change induced by devaluation would be similar for the devalued and the valued cues. A negative index value of the index represents an unexpected behavior, since the decrease would be larger for the valued cue compared with the devalued cue. We estimated the model with one to five possible latent clusters and to achieve a stable solution we run each model 200 times. The algorithm iterates between computing the expectation of the log likelihood and maximizing it to find the optimal number of latent clusters that best explain the distribution of the behavioral adaptation index. We use the Bayesian Information Criterion (BIC) criterion generated by each model to select the number of latent clusters that best accounted for the data.

Factor analysis

To reduce the questionnaires subscales we ran an exploratory factorial analysis (EFA) using maximum likelihood estimation on the standardized subscales of the questionnaires (13 subscales in total: anxiety composite scale; work overload; social overload; pressure to perform; work discontent; excessive demands at work; lack of social recognition; social tensions; social isolation; chronic worrying). We used the package Psych (Revelle 2017) with an oblimin rotation. The "parallel analysis" method suggested a four factors solution to our data. We derived the factors loadings using a regression method. The validity coefficient (R^2 = 0.97, 0.92, 0.91, and

0.89) assessing the potential impact of factor score indeterminacy (Grice 2001) was satisfactory for deriving the scores from the EFA.

For the factor labeling, we labeled the first factor "Stress Work," since the higher loadings were related to high excessive demands at work and a high workload. We labeled the second factor "Impulsivity" since the higher loadings where the three subscales of the BIS questionnaires (motor impulsiveness; nonplanning impulsiveness, and attentional impulsiveness). We labeled the third factor "Stress Social" since all the higher loadings are related to social high demands (pressure to perform, social tensions, social overload) as well as lack of social positive events (lack of social recognition). We labeled the fourth factor "Stress Affect," since the higher loadings on this factor are associated with the presence of negative affective feelings associated with stress (anxiety, worries, and discontent) and the lack of affective support (social isolation).

Multilevel analysis

We performed linear mixed effects analyses on the relationship between the response rate per second during the free-operant task and the dimensional factors extracted through the factorial analysis. As fixed effects we entered (1) Phase: pre (last training session) or post (extinction test) devaluation, (2) Cue: valued or devalued, (3) Training: moderate or extensive, and (4) the factors extracted through the factorial analysis. As random effects we entered intercepts for participants as well as by-participant random slopes for the effect of the interaction between cue and phase. We entered Block (repetition per condition) and the Site of the data collection (Pasadena1, Pasadena2, Hamburg, and Tel-Aviv) as control factors. We used the lmer4 package (Bates et al. 2015) to build the models as follows:

Response rate per second ~ (Phase × Cue) × (Training × Factor1) + Block + Site + (1 + Phase × Cue + Block | Participant)

Response rate per second ~ (Phase × Cue) × (Training × Factor2) + Block + Site + (1 + Phase × Cue + Block | Participant)

Response rate per second ~ (Phase × Cue) × (Training × Factor3) + Block + Site + (1 + Phase × Cue + Block | Participant)

Response rate per second ~ (Phase × Cue) × (Training × Factor4) + Block + Site + (1 + Phase × Cue + Block | Participant)

We reported the P-values for the model using lmerTest package (Kuznetsova et al. 2015) and corrected it for the number of tests with a significance set at $\alpha = 0.012$.

Data Deposition

Data from the study and code for the experimental task and the statistical analysis are available through the GitHub repository: https://github.com/evapool/MULTILAB_HABIT.

References

Adams CD, Dickinson A. 1981. Instrumental responding following reinforcer devaluation. *Q J Exp Psychol* **33:** 109–121. doi:10.1080/14640748108400816

Alvares GA, Balleine BW, Guastella AJ. 2014. Impairments in goal-directed actions predict treatment response to cognitive-behavioral therapy in

social anxiety disorder. *PLoS One* **9:** e94778. doi:10.1371/journal.pone .0094778

Arnsten AF. 2015. Stress weakens prefrontal networks: molecular insults to higher cognition. *Nat Neuro* **18:** 1376. doi:10.1038/nn.4087

Balleine BW, Dezfouli A. 2019. Hierarchical action control: adaptive collaboration between actions and habits. *Front Psychol* **10:** 2735. doi:10 .3389/fpsyg.2019.02735

Balleine BW, Dickinson A. 1998. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology* **37:** 407–419. doi:10.1016/S0028-3908(98)00033-1

Balleine BW, O'Doherty JP. 2010. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology* **35:** 48–69. doi:10.1038/npp.2009.131

Bates D, Mächler M, Bolker B, Walker S. 2015. Fitting linear mixed-effects models using lme4. *J Stat Softw* **67:** 1–48. doi:10.18637/jss.v067.i01

Beck AT, Steer RA, Carbin MG. 1988. Psychometric properties of the Beck Depression Inventory: twenty-five years of evaluation. *Clin Psychol Rev* **8:** 77–100. doi:10.1016/0272-7358(88)90050-5

Borenstein M, Hedges LV, Higgins JPT, Rothstein HR. 2011. *Introduction to meta-analysis.* John Wiley & Sons, New York.

Brainard DH. 1997. The psychophysics toolbox. *Spat Vis* **10:** 433–436. doi:10.1163/156856897X00357

Camerer CF, Dreber A, Holzmeister F, Ho TH, Huber J, Johannesson M, Kirchler M, Nave G, Nosek BA, Pfeiffer T, et al. 2018. Evaluating the replicability of social science experiments in Nature and Science between 2010 and 2015. *Nat Hum Behav* **2:** 637–644. doi:10.1038/ s41562-018-0399-z

Charpentier CJ, Faulkner P, Pool ER, Ly V, Tollenaar MS, Kluen LM, Fransen A, Yamamori Y, Lally N, Mkrtchian A, et al. 2020. How representative are neuroimaging samples? Large-scale evidence for trait anxiety differences between MRI and behaviour-only research participants. *Soc Cogn Affect Neurosci* **16:** 1057–1070. doi:10.1093/scan/ nsab057

De Houwer J, Tanaka A, Moors A, Tibboel H. 2018. Kicking the habit: why evidence for habits in humans might be overestimated. *Motiv Sci* **4:** 50. doi:10.1037/mot0000065

Devilbiss DM, Spencer RC, Berridge CW. 2017. Stress degrades prefrontal cortex neuronal coding of goal-directed behavior. *Cereb Cortex* **27:** 2970–2983. doi:10.1093/cercor/bhw140

de Wit S, Kindt M, Knot SL, Verhoeven AAC, Robbins TW, Gasull-Camos J, Evans M, Mirza H, Gillan CM. 2018. Shifting the balance between goals and habits: five failures in experimental habit induction. *J Exp Psychol Gen* **147:** 1043. doi:10.1037/xge0000402

Dias-Ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Costa RM, Sousa N. 2009. Chronic stress causes frontostriatal reorganization and affects decision-making. *Science* **325:** 621–625. doi:10.1126/science.1171203

Dickinson A. 1985. Actions and habits: the development of behavioural autonomy. *Philos Trans R Soc Lond B Biol Sci* **308:** 67–78. doi:10.1098/ rstb.1985.0010

Dickinson A, Nicholas DJ, Adams CD. 1983. The effect of the instrumental training contingency on susceptibility to reinforcer devaluation. *Q J Exp Psychol* **35:** 35–51. doi:10.1080/14640748308400912

Dickinson A, Balleine B, Watt A, Gonzalez F, Boakes RA. 1995. Motivational control after extended instrumental training. *Anim Learn Behav* **23:** 197–206. doi:10.3758/BF03199935

Ersche KD, Lim TV, Ward LH, Robbins TW, Stochl J. 2017. Creature of habit: a self-report measure of habitual routines and automatic tendencies in everyday life. *Pers Individ Diff* **116:** 73–85. doi:10.1016/j.paid.2017.04 .024

Everitt BJ, Robbins TW. 2016. Drug addiction: updating actions to habits to compulsions 10 years on. *Annu Rev Psychol* **67:** 23–50. doi:10.1146/ annurev-psych-122414-033457

Fernando A, Urcelay G, Mar A, Dickinson A, Robbins TW. 2014. Free-operant avoidance behavior by rats after reinforcer revaluation using opioid agonists and D-amphetamine. *J Neurosci* **34:** 6286–6293. doi:10.1523/ JNEUROSCI.4146-13.2014

Foa EB, Huppert JD, Leiberg S, Langner R, Kichic R, Hajcak G, Salkovskis PM. 2002. The obsessive-compulsive inventory: development and validation of a short version. *Psychol Assess* **14:** 485. doi:10.1037/1040-3590.14.4 .485

Gillan CM, Kosinski M, Whelan R, Phelps EA, Daw ND. 2016. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *Elife* **5:** e11305. doi:10.7554/eLife.11305

Gillan CM, Vaghi MM, Hezemans FH, van Ghesel Grothe S, Dafflon J, Brühl AB, Savulich G, Robbins TW. 2021. Experimentally induced and real-world anxiety have no demonstrable effect on goal-directed behaviour. *Psychol Med* **51:** 1467–1478. doi:10.1017/S0033291 720000203

Goldfarb EV. 2019. Enhancing memory with stress: progress, challenges, and opportunities. *Brain Cogn* **133:** 94–105. doi:10.1016/j.bandc.2018 .11.009

Goldfarb EV. 2020. Participant stress in the COVID-19 era and beyond. *Nat Rev Neurosci* **21:** 663–664. doi:10.1038/s41583-020-00388-7

Goldfarb EV, Shields GS, Daw ND, Slavich GM, Phelps EA. 2017. Low lifetime stress exposure is associated with reduced stimulus–response memory. *Learn Mem* **24:** 162–168. doi:10.1101/lm.045179.117

Gremel C, Costa R. 2013. Premotor cortex is critical for goal-directed actions. *Front Comput Neurosci* **7:** 110. doi:10.3389/fncom.2013.00110

Grice JW. 2001. Computing and evaluating factor scores. *Psychol Methods* **6:** 430. doi:10.1037/1082-989X.6.4.430

Hardwick RM, Forrence AD, Krakauer JW, Haith AM. 2019. Time-dependent competition between goal-directed and habitual response preparation. *Nat Hum Behav* **3:** 1252–1262. doi:10.1038/s41562-019-0725-0

Hartogsveld B, van Ruitenbeek P, Quaedflieg CWEM, Smeets T. 2020. Balancing between goal-directed and habitual responding following acute stress. *Exp Psychol* **67:** 99–111. doi:10.1027/1618-3169/a000485

Hilário MR, Clouse E, Yin HH, Costa RM. 2007. Endocannabinoid signaling is critical for habit formation. *Front Int Neurosci* **1:** 6.

Hogarth L. 2020. Addiction is driven by excessive goal-directed drug choice under negative affect. *Neuropharmacology* **5:** 720–735.

Holland PC. 2004. Relations between Pavlovian-instrumental transfer and reinforcer devaluation. *J Exp Psychol Anim Behav Process* **30:** 104. doi:10 .1037/0097-7403.30.2.104

Huys QJM, Maia TV, Frank MJ. 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. *Nat Neurosci* **19:** 404. doi:10 .1038/nn.4238

Kuznetsova A, Brockhoff PB, Christensen RHB. 2015. lmerTest: tests in linear mixed effects models. R package version 2 (0).

Lazarus RS, Folkman S. 1984. *Stress, appraisal, and coping.* Springer Publishing Company, New York.

Leisch F. 2004. Flexmix: a general framework for finite mixture models and latent glass regression in R. *J Stat Softw* **11:** 1–18. doi:10.18637/jss .v011.i08

Lupien SJ, Maheu F, Tu M, Fiocco A, Schramek TE. 2007. The effects of stress and stress hormones on human cognition: implications for the field of brain and cognition. *Brain Cogn* **65:** 209–237. doi:10.1016/j.bandc.2007 .02.007

Luque D, Molinero S, Watson P, López FJ, Le Pelley ME. 2019. Measuring habit formation through goal-directed response switching. *J Exp Psychol Gen* **149:** 1449–1459. doi:10.1037/xge0000722

Margittai Z, Gideon N, Strombach T, van Wingerden M, Schwabe L, Kalenscher T. 2016. Exogenous cortisol causes a shift from deliberative to intuitive thinking. *Psychoneuroendocrinology* **64:** 131–135. doi:10 .1016/j.psyneuen.2015.11.018

Meier JK, Staresina BP, Schwabe L. 2021. Stress diminishes outcome but enhances response representations during instrumental learning. bioRxiv doi:10.1101/2021.02.12.430935

Moors A, Boddez Y, Houwer JD. 2017. The power of goal-directed processes in the causation of emotional and other actions. *Emot Rev* **9:** 310–318. doi:10.1177/1754073916669595

Morris SB, DeShon RP. 2002. Combining effect size estimates in meta-analysis with repeated measures and independent-groups designs. *Psychol Methods* **7:** 105. doi:10.1037/1082-989X.7.1.105

Nave G, Daviet R, Nadler R, Zava D, Camerer C. 2020. Reflecting on the evidence: a reply to Knight, McShane, et al. (2020). *Psychol Sci* **31:** 898–900. doi:10.1177/0956797620930966

Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND. 2013. Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci* **110:** 20941–20946. doi:10.1073/pnas.1312011110

Ouellette JA, Wood W. 1998. Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychol Bull* **124:** 54. doi:10.1037/0033-2909.124.1.54

Packard MG. 1999. Glutamate infused posttraining into the hippocampus or caudate-putamen differentially strengthens place and response learning. *Proc Natl Acad Sci* **96:** 12881–12886. doi:10.1073/pnas.96.22 .12881

Patton JH, Stanford MS, Barratt ES. 1995. Factor structure of the Barratt impulsiveness scale. *J Clin Psychol* **51:** 768–774. doi:10.1002/1097-4679 (199511)51:6<768::AID-JCLP2270510607>3.0.CO;2-1

Patzelt EH, Kool W, Millner AJ, Gershman SJ. 2019. Incentives boost model-based control across a range of severity on several psychiatric constructs. *Biol Psychiatry* **85:** 425–433. doi:10.1016/j.biopsych.2018.06 .018

Pelli DG. 1997. The VideoToolbox software for visual psychophysics: transforming numbers into movies. *Spat Vis* **10:** 437–442. doi:10.1163/ 156856897X00366

Perez OD, Dickinson A. 2020. A theory of actions and habits: the interaction of rate correlation and contiguity systems in free-operant behavior. *Psychol Rev* **127:** 945–971. doi:10.1037/rev0000201

Perez OD, Soto FA. 2020. Evidence for a dissociation between causal beliefs and instrumental actions. *Q J Exp Psychol* **73:** 495–503. doi:10.1177/ 1747021819899808

Petrowski K, Paul S, Albani C, Brähler E. 2012. Factor structure and psychometric properties of the Trier Inventory for Chronic Stress (TICS) in a representative German sample. *BMC Med Res Methodol* **12:** 42. doi:10 .1186/1471-2288-12-42

Pool ER, Sander D. 2019. Vulnerability to relapse under stress: insights from affective neuroscience. *Swiss Med Wkly* **149:** 4748. doi:10.4414/smw .2019.20151

Quaedflieg CWEM, Stoffregen H, Sebalo I, Smeets T. 2019. Stress-induced impairment in goal-directed instrumental behaviour is moderated by baseline working memory. *Neurobiol Learn Mem* **158:** 42–49. doi:10 .1016/j.nlm.2019.01.010

Radenbach C, Reiter AMF, Engert V, Sjoerds Z, Villringer A, Heinze HJ, Deserno L, Schlagenhauf F. 2015. The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology* **53:** 268–280. doi:10.1016/j.psyneuen.2014.12.017

Reed P. 2001. Human schedule performance with hypothetical monetary reinforcement. *Eur J Behav Anal* **2:** 225–234. doi:10.1080/15021149 .2001.11434197

Revelle WR. 2017. Psych: procedures for personality and psychological research. R package version 1.0-95.

Rouder JN, Morey RD, Speckman PL, Province JM. 2012. Default Bayes factors for ANOVA designs. *J Math Psychol* **56:** 356–374. doi:10.1016/j .jmp.2012.08.001

Scherer KR. 2005. What are emotions? And how can they be measured? *Soc Sci Inf* **44:** 695–729. doi:10.1177/0539018405058216

Schwabe L, Wolf OT. 2009. Stress prompts habit behavior in humans. *J Neurosci* **29:** 7191–7198. doi:10.1523/JNEUROSCI.0979-09.2009

Schwabe L, Wolf OT. 2010. Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. *Psychoneuroendocrinology* **35:** 977–986. doi:10.1016/j.psyneuen .2009.12.010

Schwabe L, Dalm S, Schachinger H, Oitzl MS. 2008. Chronic stress modulates the use of spatial and stimulus-response learning strategies in mice and man. *Neurobiol Learn Mem* **90:** 495–503. doi:10.1016/j.nlm .2008.07.015

Schwabe L, Höffken O, Tegenhoff M, Wolf OT. 2011. Preventing the stress-induced shift from goal-directed to habit action with a β-adrenergic antagonist. *J Neurosci* **37:** 17317–17325. doi:10.1523/ JNEUROSCI.3304-11.2011

Simons DJ, Shoda Y, Lindsay SD. 2017. Constraints on generality (COG): a proposed addition to all empirical papers. *Perspect Psychol Sci* **12:** 1123– 1128. doi:10.1177/1745691617708630

Singmann H, Bolker B, Westfall J, Aust F. 2015. Afex: analysis of factorial experiments. R package version 0.13–145.

Sjoerds Z, de Wit S, van den Brink W, Robbins TW, Beekman ATF, Penninx BWJH, Veltman DJ. 2013. Behavioral and neuroimaging evidence for overreliance on habit learning in alcohol-dependent patients. *Transl Psychiatry* **3:** e337. doi:10.1038/tp.2013.107

Snorrason I, Lee HJ, de Wit S, Woods DW. 2016. Are nonclinical obsessive-compulsive symptoms associated with bias toward habits? *Psychiatry Res* **241:** 221–223. doi:10.1016/j.psychres.2016.04.067

Soares JM, Sampaio A, Ferreira LM, Santos NC, Marques F, Palha JA, Cerqueira JJ, Sousa N. 2012. Stress-induced changes in human decision-making are reversible. *Transl Psychiatry* **2:** e131. doi:10.1038/tp .2012.59

Spielberger CD. 2013. *Anxiety: current trends in theory and research.* Academic Press, New York.

Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA. 1983. *Manual for the State-Trait Anxiety Inventory.* Consulting Psychologists Press, Palo Alto, CA.

Starcke K, Brand M. 2012. Decision making under stress: a selective review. *Neurosci Biobehav Rev* **36:** 1228–1248. doi:10.1016/j.neubiorev.2012.02 .003

Tricomi E, Balleine BW, O'Doherty JP. 2009. A specific role for posterior dorsolateral striatum in human habit learning. *Eur J Neurosci* **29:** 2225– 2232. doi:10.1111/j.1460-9568.2009.06796.x

Viechtbauer W. 2010. Conducting meta-analyses in R with the meta for package. *J Statist Softw* **36:** 1–48. doi:10.18637/jss.v036.i03

Voon V, Derbyshire K, Rück C, Irvine MA, Worbe Y, Enander J, Schreiber LRN, Gillan C, Fineberg NA, Sahakian BJ. 2015. Disorders of compulsivity: a common bias towards learning habits. *Mol Psychiatry* **20:** 345–352. doi:10.1038/mp.2014.44

Whiteside SP, Lynam DR. 2001. The five factor model and impulsivity: using a structural model of personality to understand impulsivity. *Pers Individ Dif* **30:** 669–689. doi:10.1016/S0191-8869(00)00064-7

Wiltgen BJ, Sinclair C, Lane C, Barrows F, Molina M, Chabanon-Hicks C. 2012. The effect of ratio and interval training on Pavlovian-instrumental transfer in mice. *PLoS One* **7:** e48227. doi:10.1371/journal.pone .0048227