

Don't Do It Again! Directed Forgetting of Habits

Psychological Science 2014, Vol. 25(6) 1242–1248 © The Author(s) 2014 Reprints and permissions: sagepub.com/journalsPermissions.nav DOI: 10.1177/0956797614526063 pss.sagepub.com



Gesine Dreisbach and Karl-Heinz T. Bäuml

Regensburg University

Abstract

Most daily routines are determined by habits. However, the experienced ease and automaticity of habit formation and execution come at a cost when habits that are no longer appropriate must be overcome. So far, proactive and reactive control strategies that prevent inappropriate habit execution either by preparation or "on the fly" have been identified. Here, we present evidence for a third, retroactive control strategy. In two experiments using the list method of directed forgetting, the accessibility of newly learned and practiced stimulus-response rules was significantly reduced when participants were cued to forget the rules rather than to remember them. The results thus show that directed forgetting, so far observed and investigated only for episodic memory traces, can also be applied to habits. The findings further emphasize the adaptive value of forgetting and can be taken as evidence of a retroactive strategy of habit control.

Keywords

habit, cognitive control, directed forgetting, response compatibility

Received 9/17/13; Revision accepted 2/7/14

Human beings are creatures of habit. Many daily routines, such as getting up in the morning, taking a shower, and driving to work, take place without much thought and deliberate intention. But at the same time, people are not slaves to their habits and are clearly able to override habitual response tendencies in favor of more appropriate behavior in a given situation (Miller & Cohen, 2001). For example, if you normally drink beer directly from the bottle, you are able to refrain from this habit and use a glass instead when in a fancy restaurant. This ability to resist routines and habits and to shield goals and behavior from strong but currently inappropriate response tendencies is usually referred to as cognitive control. Reactive control strategies enable the context-sensitive adjustment of cognitive control in response to experienced conflicts or task demands, and proactive control strategies support the adjustment of cognitive control in anticipation of expected task demands. Previous work has provided evidence for both reactive control strategies (e.g., Botvinick, Braver, Barch, Carter, & Cohen, 2001; Gratton, Coles, & Donchin, 1992; Kerns et al., 2004) and proactive control strategies (e.g., Braver, 2012; Dreisbach & Haider, 2006; Gollwitzer, 1999). Here, we present evidence for a third strategy, retroactive control, that uses deliberate forgetting to prevent interference from habits.

One way of directly investigating resistance against habitual behavior in the laboratory (introduced by Ach, 1910/2006) is to first establish habits by having participants practice an arbitrary stimulus-response (SR) rule, and then change the rule and measure the resulting performance cost. There is broad consensus that repeated SR associations leave traces in long-term memory, such that the occurrence of a given stimulus eventually automatically retrieves the associated response (e.g., Hommel, 2000; Logan, 1988). A single SR episode, or even merely receiving instructions about or planning an SR rule, can trigger the automatic retrieval of the entire SR association when the stimulus is encountered (e.g., Hommel, 1998; Wenke, Gaschler, & Nattkemper, 2007). A common precondition in all these cases is that at some point during the experiment, participants must form the intention to react to the stimulus with a prespecified response (Hommel, 2000), a precondition that also holds for many everyday habits. We address here whether this intention and the consequential automatic retrieval can be undone by directed forgetting.

Corresponding Author:

Gesine Dreisbach, Department of Experimental Psychology, University of Regensburg, Universitätsstraße 31, D-93053 Regensburg, Germany E-mail: gesine.dreisbach@psychologie.uni-regensburg.de

Table 1. Example of the Stimulus-Response Mappings in Experiment 1

Block	Left response	Right response
Block 1	Wolke ("cloud," f)	Glocke ("bell," f)
	Finger ("finger," m)	Stuhl ("chair," m)
Block 2	Trommel ("drum," f)	Maske ("mask," f)
	Hafen ("haven," m)	Keller ("basement," m)
Block 3	Wolke (f)	Finger (m)
	Glocke (f)	Stuhl (m)
	Trommel (f)	Hafen (m)
	$\mathbf{Maske}(f)$	Keller (m)

Note: In this example, in Block 3, words with feminine grammatical gender (f) were mapped to the left key, and words with masculine gender (m) were mapped to the right key. On compatible trials, the mapping between stimulus and response did not change from Block 1 or 2 to Block 3; on incompatible trials (indicated in boldface), the mapping changed.

In the study of episodic memory, *directed forgetting* refers to a paradigm in which people intentionally forget previously studied material when cued to do so (e.g., Bjork, 1970). In the list method of this paradigm, participants study a list of items and then receive a cue either to forget or to continue remembering the list. After subsequent study of another list, participants are asked to recall the items on the first list. Typically, the forget cue impairs recall of the items, relative to the remember cue, a finding attributed to impaired access to the original encoding context (Geiselman, Bjork, & Fishman, 1983; Sahakyan & Kelley, 2002).

Applying the logic of directed-forgetting experiments to habits, we asked participants to learn arbitrary right/left responses to a small set of words: four words in Block 1, four different words in Block 2, and the eight words from Blocks 1 and 2 in Block 3. Between Blocks 1 and 2, half of the participants were instructed to forget the Block 1 associations. In Block 3, four of the associations were the same as they had been in the earlier block (compatible trials), whereas the other four were the opposite of what they had been earlier (incompatible trials; Table 1). The compatibility effect (i.e., the performance difference between compatible and incompatible trials) then served as a measure of the persisting activation of the competing (i.e., original) SR rules. If directed forgetting works for habits just as it works for episodic memory traces in general, the compatibility effect for Block 1 associations should be smaller after a forget instruction compared with a remember instruction (or might even be absent).

Experiment 1

Method

Participants. Forty students from Regensburg University (21 female, 19 male; mean age = 20.7 years, SD = 3.7

years) participated for partial course credit or €4. They provided informed consent and were debriefed after the session. Half the participants were randomly assigned to the remember condition, and the other half to the forget condition.

Stimuli and procedure. Four German words served as stimuli in Block 1, and four different German words were stimuli in Block 2. Each set of four words contained two words of masculine grammatical gender and two of feminine grammatical gender (for a list of the words, see Table 1). Participants responded to each word by pressing the "Y" key (left) or the "M" key (right), using a computer keyboard with the QWERTZ layout. In each block, one word of each gender was mapped to the left response, and one word of each gender was mapped to the right response. Within these constraints, stimulus sets and SR mappings were counterbalanced across blocks and participants.

In Block 3, the test block, all eight words were presented. Participants were asked to press the left key in response to a masculine word and the right key in response to a feminine word, or vice versa; the left/right positions of the response keys for masculine and feminine words were counterbalanced across participants. Thus, half of the SR mappings in Block 3 were compatible with the mappings from the previous blocks, and the other half were incompatible (see Table 1).

Each trial started with a fixation cross (250 ms), which was followed by a word stimulus that remained on the screen until the participant responded. Feedback was given only for errors (in the form of the German word "Fehler"). The intertrial interval was 1,000 ms after a correct response and 2,000 ms following an error. Words were presented in random order. Each word in Block 1 was presented 18 times, each word in Block 2 was presented 20 times, and each word in Block 3 was presented 14 times. Instructions emphasized speed and accuracy equally.

In the forget condition, the first block ended with a simulated computer crash (indicated by a blue screen plus an error message). The experimenter then stepped in, apologized for having started the wrong program, and kindly asked the participant to start over again with the correct program. The participant was then casually told to try to forget what he or she had just done (i.e., to forget the words and SR mappings in Block 1). The experimenter then started the new program (i.e., Blocks 2 and 3; see also Dobler & Bäuml, 2012).

In the remember condition, Block 1 ended with a screen telling participants to let the experimenter know that Block 1 was complete. The experimenter then casually informed the participants that they should remember what they had just done. The experimenter then started the new program (Blocks 2 and 3).

1244 Dreisbach, Bäuml

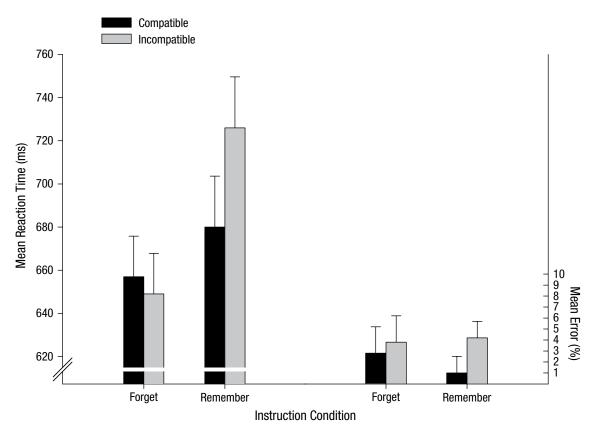


Fig. 1. Experiment 1 results: mean reaction time (left) and error rate (right) in Block 3 as a function of compatibility and instruction condition. Error bars represent 95% within-subjects confidence intervals (Loftus & Masson, 1994) derived from the corresponding compatibility effect.

After Block 2, participants in both conditions were informed that on the following trials, they should decide whether a given word was of masculine or feminine grammatical gender. In a postexperimental interview, none of the participants in the forget condition expressed any suspicion about the computer crash.

Design. In Block 3, a 2 (compatibility: compatible, incompatible) × 2 (instruction: forget, remember) mixed-factors design was applied. Compatibility was manipulated within participants, and instruction was manipulated between participants. Reaction times (RTs) and error rates for Block 1 items served as dependent measures.

Results and discussion

Trials responded to incorrectly and trials with RTs more than 3 standard deviations from the individual cell mean (1.65% of all trials) were excluded from RT analyses. To examine the prediction that the compatibility effect for words from Block 1 should be smaller in the forget than in the remember condition, we ran a 2 (compatibility:

compatible, incompatible) × 2 (instruction: forget, remember) mixed-factors analysis of variance (ANOVA) on RTs for responses to Block 1 words (see Fig. 1). Analysis revealed a main effect of instruction, F(1, 38) = 4.8, p < .05, $\eta_p^2 = .11$. RTs were generally slower in the remember condition (M = 703 ms) than in the forget condition (M =653 ms). Compatibility had a marginally significant effect, F(1, 38) = 3.1, p = .08. Most important, the interaction of instruction and compatibility was significant, F(1, 38) =6.1, p < .02, $\eta_p^2 = .14$. RTs on compatible and incompatible trials did not differ in the forget condition, F < 1, whereas responses on compatible trials (M = 680 ms) were significantly faster than responses on incompatible trials (M =726 ms) in the remember condition, $F(1, 38) = 8.9, p < .01.^{2}$ (For practice effects within Block 3, see Fig. S1 and Table S1 in the Supplemental Material available online.)

An analogous 2×2 ANOVA was run on the error rates for Block 1 words (Fig. 1). Only the main effect of compatibility reached significance, F(1, 38) = 4.2, p < .05, $\eta_p^2 = .10$; participants generally made more errors on incompatible trials (4%) than on compatible trials (1.9%). Although the interaction between compatibility and

instruction was not statistically significant, F = 1.0, p = .3, error rates were significantly greater for incompatible than for compatible trials in the remember condition, F(1, 38) = 4.75, p < .05, $\eta_p^2 = .11$, but not in the forget condition, F < 1, p > .4.

Experiment 1 provides the first evidence that newly learned and practiced SR rules can be subject to directed forgetting. Using the compatibility effect for Block 1 words as a measure of persisting activation of previously learned SR rules, we obtained clear evidence that the forget instruction reduced the impact of these SR associations: The compatibility effect was significantly reduced in the forget condition compared with the remember condition. Indeed, the compatibility effect was virtually absent in the forget condition.

In Experiment 1, we used arbitrary SR rules in Blocks 1 and 2, but a task rule in the test phase. There is much evidence that task rules, which allow for natural categorization (here, masculine vs. feminine gender), are easily formed and more persistent than arbitrary SR rules (Dreisbach, 2012; Dreisbach, Goschke, & Haider, 2006, 2007). Our design may have made it rather easy for us to find directed forgetting of habits, and it was unclear whether the results would generalize to task rules based on a natural category and long-term associations between stimulus and response.

Therefore, in Experiment 2, we tested the robustness of the directed-forgetting effect using task rules in Blocks 1 and 2. In Block 1, participants responded to small numbers by pressing the left key and to large numbers by pressing the right key. This task exploited the natural tendency to associate larger numbers with spatial positions more to the right and smaller numbers with spatial positions more to the left—the SNARC (spatial-numerical association of response codes) effect (Dehaene, Bossini, & Giraux, 1993). In Block 2, the stimuli were letters that had to be categorized as consonants or vowels. In Block 3, participants categorized the numbers from Block 1 according to parity (i.e., whether they were odd or even). This rule change again made half of the stimuli compatible and the other half incompatible. We reasoned that the compatibility effect might be harder to reduce in this experiment than in Experiment 1, because the SR rule in question relied not only on recent practice but also on the natural tendency to spatially associate number magnitude. A significant reduction of the compatibility effect in the forget condition would thus provide even more compelling evidence that automatic retrieval of habits can be hampered by intention.

Experiment 2

Method

Participants. Eighty students³ (60 female, 20 male; mean age = 22.3 years, SD = 2.7 years) from Regensburg

University participated for partial course credit or €4. Two participants were excluded because of high error rates, and 2 participants in the forget condition were excluded because they expressed suspicion about the computer crash. They were replaced by another 4 participants.

Stimuli and procedure. The numbers 1 to 4 and 6 to 9 served as stimuli in Blocks 1 and 3. In Block 2, the letters "A," "E," "O," "U," "C," "F," "S," and "R" were presented. In Block 1, participants judged whether each number was less than 5 (left key) or greater than 5 (right key). In Block 2, participants decided whether each letter was a vowel (left key) or consonant (right key). Finally, in Block 3, only numbers from Block 1 were presented (to avoid task switching within the block), and participants were asked to decide whether each number was odd (left key) or even (right key). In each block, the eight stimuli were presented 10 times each, so block length was constant at 80 trials. All other procedural details, including the directed-forgetting manipulation, followed the procedure of Experiment 1.

Results and discussion

Block 3 trials responded to incorrectly and trials with RTs more than 3 standard deviations from the individual cell mean (1.48% of all trials) were again excluded from RT analysis. Mean RTs for Block 1 items were entered into a 2 (instruction; between participants) \times 2 (compatibility; within participants) mixed-factors ANOVA (see Fig. 2). A significant main effect of compatibility, F(1, 78) = 33.86, p < .01, $\eta_p^2 = .30$, was qualified by a significant interaction with instruction, F(1, 78) = 5.1, p < .03, $\eta_p^2 = .06$. RTs were generally faster on compatible trials than on incompatible trials (M = 580 ms vs. 613 ms). More important, this compatibility effect was smaller in the forget condition than in the remember condition (M = 20 ms vs. 46ms). Compatibility effects were significant for both the forget condition, F(1, 78) = 6.36, p < .05, $\eta_p^2 = .07$, and the remember condition, F(1, 78) = 32.57, p < .01, $\eta_p^2 =$.29 (for practice effects within Block 3, see Fig. S2 and Table S2 in the Supplemental Material).

Analysis of the error data again yielded a significant main effect of compatibility, F(1, 78) = 19.3, p < .01, $\eta_p^2 = .19$; participants made fewer errors on compatible trials (2% of all trials) than on incompatible trials (5.0%). The main effect of instruction approached significance, F(1, 78) = 2.9, p = .08. Again, compatibility effects were significant in both the forget condition, F(1, 78) = 4.69, p < .05, $\eta_p^2 = .06$, and the remember condition, F(1, 78) = 16.33, p < .01, $\eta_p^2 = .17$.

The results of Experiment 2 confirm the results of Experiment 1. The compatibility effect was again significantly smaller in the forget condition than in the

1246 Dreisbach, Bäuml

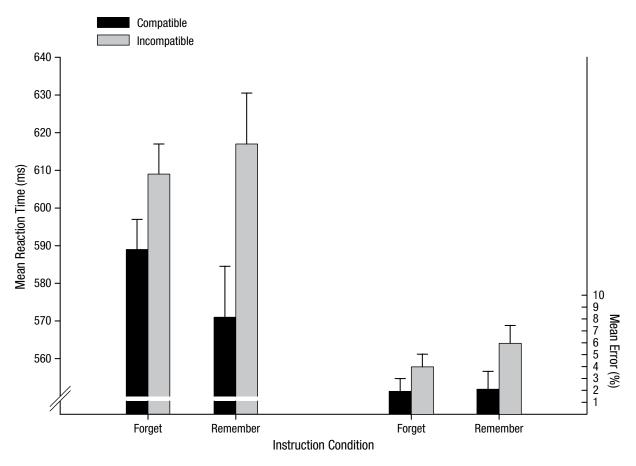


Fig. 2. Experiment 2 results: mean reaction time (left) and error rate (right) in Block 3 as a function of compatibility and instruction condition. Error bars represent 95% within-subjects confidence intervals (Loftus & Masson, 1994) derived from the corresponding compatibility effect.

remember condition, which indicates that the forget instruction reduced the influence of SR associations acquired in Block 1.

General Discussion

The present results demonstrate for the first time that the influence of newly established habits can be reduced by means of directed forgetting. We used compatibility effects as a measure of the persisting activation of formerly established SR associations. For the remember condition, we found significant compatibility effects, but for the forget condition, we found reduced effects (for the more natural task rule in Experiment 2) or even no effect (for the arbitrary SR rules in Experiment 1). These findings indicate that automatic retrieval of practiced SR associations can be prevented when people are cued to do so.

Our results are consistent with the literature on both habits and directed forgetting. Because memory for habits seems to depend on access to the encoding context (Wood & Neal, 2007; Wood, Tam, & Guerrero Witt,

2005), and the forget cue supposedly impairs access to the original encoding context (Geiselman et al., 1983; Sahakyan & Kelley, 2002), memory for habits should be subject to directed forgetting, which is what our results show. Moreover, the stronger directed-forgetting effect in Experiment 1 than in Experiment 2 suggests that accessibility of long-term associations is not as easily reduced when the associations are based on a natural categorization rule as when they are based on arbitrary SR rules. The fact that the compatibility effect for Block 1 stimuli in the forget condition was still significant (though reduced) in Experiment 2 but not in Experiment 1 also supports this assumption. Although comparisons across experiments should be treated with caution, this finding might be related to the observation that strong habits are less dependent on encoding context than are weak habits (Wood et al., 2005) and further confirms the greater persistence of general task rules compared with SR rules (Dreisbach, 2012; Mayr & Bryck, 2005).

Habits rely heavily on context: They are cued by recurring context features and are less accessible when the context changes (Wood & Neal, 2007). By contrast,

implicit memory for encounters with single words is unaffected by context change and is equivalent in samecontext and different-context conditions (Mulligan, 2011; Parker, Dagnall, & Coyle, 2007). Therefore, our results do not contradict previous research in which tests of implicit memory revealed no directed forgetting (e.g., Basden, Basden, & Gargano, 1993). The finding of directed forgetting for habits is also not in conflict with studies showing that retrieval of a subset of List 1 items can eliminate directed forgetting for the other list items, probably by reactivating the original encoding context (Bäuml & Samenieh, 2010, 2012). In the present study, stimulus and response features of Block 1 were also present in Block 3, but both the change in response criteria from Block 1 to Block 3 and the presence of new associations in Block 3 should have prevented contextual drift back to Block 1.

In conclusion, there seem to be at least three control strategies available to secure appropriate behavior in our initial beer-drinking example. First, context-sensitive reactive control may detect the conflict between bottle and glass, and ideally bias the response "just in time" (Botvinick et al., 2001; Braver, 2012). Second, proactive control, in the form of an implementation intention (e.g., "If I am in a fancy restaurant, then I will pour the beer into the glass before I take a sip") may prevent the response conflict before it even arises (Gollwitzer, 1999). And finally, retroactive control (e.g., "I will never drink beer from the bottle again"), as discussed here, may generally reduce accessibility of the inappropriate habit. Although the precise mechanisms underlying these control processes await future investigation, the present results clearly indicate that the cognitive system is well equipped to adapt behavior by biasing habits in accordance with current and future goals and intentions.

Author Contributions

G. Dreisbach developed the study concept. Both authors contributed to the study design. The experiment was programmed by G. Dreisbach. Testing and data collection was supervised by G. Dreisbach. G. Dreisbach performed the data analysis and drafted the manuscript, and K.-H. T. Bäuml provided critical revisions. Both authors provided substantial revisions in response to the reviewers and approved the final version of the manuscript for submission.

Declaration of Conflicting Interests

The authors declared that they had no conflicts of interest with respect to their authorship or the publication of this article.

Supplemental Material

Additional supporting information may be found at http://pss.sagepub.com/content/by/supplemental-data

Notes

- 1. In directed forgetting, study of a postcue (second) list is necessary for the forget cue to induce forgetting (e.g., Pastötter & Bäuml, 2007). Therefore, we incorporated a postcue block of associations before the final test in Block 3.
- 2. There was a significant compatibility effect for Block 2 words, F(1, 38) = 19.6, p < .01, that did not differ between the two instruction conditions, F < 1.
- 3. Sample size in Experiment 1 was determined on the basis of prior directed-forgetting studies (e.g., Bäuml & Samenieh, 2010, 2012). Given the medium effect size obtained in Experiment 1, we doubled the sample size in Experiment 2 to increase power.

References

- Ach, N. (2006). *On volition* (T. Herz, Trans.). Retrieved from http://www.uni-konstanz.de/kogpsych/ach.htm (Original work published 1910)
- Basden, B. H., Basden, D. R., & Gargano, G. J. (1993). Directed forgetting in implicit and explicit memory tests: A comparison of methods. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 603–616.
- Bäuml, K.-H. T., & Samenieh, A. (2010). The two faces of memory retrieval. *Psychological Science*, *21*, 793–795.
- Bäuml, K.-H. T., & Samenieh, A. (2012). Selective memory retrieval can impair and improve retrieval of other memories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 38, 488–494.
- Bjork, R. A. (1970). Positive forgetting: The noninterference of items intentionally forgotten. *Journal of Verbal Learning and Verbal Behavior*, 9, 255–268.
- Botvinick, M. W., Braver, T. S., Barch, D. M., Carter, C. S., & Cohen, J. D. (2001). Conflict monitoring and cognitive control. *Psychological Review*, *108*, 624–652.
- Braver, T. S. (2012). The variable nature of cognitive control: A dual mechanisms framework. *Trends in Cognitive Sciences*, *16*, 106–113.
- Dehaene, S., Bossini, S., & Giraux, P. (1993). The mental representation of parity and numerical magnitude. *Journal of Experimental Psychology: General*, 122, 371–396.
- Dobler, I. M., & Bäuml, K.-H. T. (2012). Dissociating the two faces of selective memory retrieval. *Memory*, 20, 478–486.
- Dreisbach, G. (2012). Mechanisms of cognitive control: The functional role of task rules. *Current Directions in Psychological Science*, 21, 227–231.
- Dreisbach, G., Goschke, T., & Haider, H. (2006). Implicit task sets in task switching? *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *32*, 1221–1233.
- Dreisbach, G., Goschke, T., & Haider, H. (2007). The role of task rules and stimulus-response mappings in the task switching paradigm. *Psychological Research*, 71, 383–392.
- Dreisbach, G., & Haider, H. (2006). Preparatory adjustment of cognitive control in the task switching paradigm. *Psychonomic Bulletin & Review*, *13*, 334–338.
- Geiselman, R. E., Bjork, R. A., & Fishman, D. L. (1983). Disrupted retrieval in directed forgetting: A link with post-hypnotic amnesia. *Journal of Experimental Psychology: General*, 112, 58–72.

1248 Dreisbach, Bäuml

Gollwitzer, P. M. (1999). Implementation intentions: Strong effects of simple plans. American Psychologist, 54, 493–503.

- Gratton, G., Coles, M. G., & Donchin, E. (1992). Optimizing the use of information: Strategic control of activation of responses. *Journal of Experimental Psychology: General*, 121, 480–506.
- Hommel, B. (1998). Event files: Evidence for automatic integration of stimulus-response episodes. Visual Cognition, 5, 183–216.
- Hommel, B. (2000). The prepared reflex: Automaticity and control in stimulus-response translation. In S. Monsell & J. Driver (Eds.), Attention and performance XVIII: Control of cognitive processes (pp. 247–273). Cambridge, MA: MIT Press.
- Kerns, J. G., Cohen, J. D., MacDonald, A. W., III, Cho, R. Y., Stenger, V. A., & Carter, C. S. (2004). Anterior cingulate conflict monitoring and adjustments in control. *Science*, 303, 1023–1026.
- Loftus, G. R., & Masson, M. E. J. (1994). Using confidence intervals in within-subjects designs. *Psychonomic Bulletin & Review*, 1, 476–490.
- Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review*, *95*, 492–527.
- Mayr, U., & Bryck, R. L. (2005). Sticky rules: Integration between abstract rules and specific actions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 337–350.

- Miller, E. K., & Cohen, J. D. (2001). Integrative theory of PFC function. *Annual Review of Neuroscience*, 24, 167–202.
- Mulligan, N. W. (2011). Conceptual implicit memory and environmental context. Consciousness and Cognition, 20, 737–744.
- Parker, A., Dagnall, N. A., & Coyle, A. (2007). Environmental context effects in conceptual explicit and implicit memory. *Memory*, 15, 423–434.
- Pastötter, B., & Bäuml, K.-H. (2007). The crucial role of postcue encoding in directed forgetting and context-dependent forgetting. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 33,* 977–982.
- Sahakyan, L., & Kelley, C. M. (2002). A contextual change account of the directed forgetting effect. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 28, 1064– 1072.
- Wenke, D., Gaschler, R., & Nattkemper, D. (2007). Instructioninduced feature binding. *Psychological Research*, 71, 92–106.
- Wood, W., & Neal, D. T. (2007). A new look at habits and the habit–goal interface. *Psychological Review*, 4, 843–863.
- Wood, W., Tam, L., & Guerrero Witt, M. (2005). Changing circumstances, disrupting habits. *Journal of Personality and Social Psychology*, 88, 918–933.