The Contribution of Falsework Fading to the Stimulus Generalization of a Skill

Sharon Bertsch and Robert J. Zeglin
University of Pittsburgh at Johnstown

It has been established in the animal literature that context generalization gradients flatten over time—performance of a learned behavior is more likely to be expressed in novel situations after longer retention intervals. We attempted to extend this finding to stimulus generalization among human participants. One hundred twenty participants were trained to trace mirror-reversed shapes using either a blocked or variable schedule and then tested on a novel shape immediately, after 2 days or after 1 week. We found that variable training led to more effective transfer performance immediately, but while those in the blocked condition improved their transfer performance over time, those in the variable training condition declined. We believe that the improvement of the blocked trained group is due to a reduction in the strength of the "falsework" connection initially formed between the training stimulus and the new skill.

In 1992, Schmidt and Bjork reviewed the literature on the training, retention and generalization of new skills. They suggested that the very attributes of practice that result in rapid acquisition are also those least likely to promote either long term retention or effective transfer to novel stimuli. One example of these attributes is how the training trials are scheduled. Consistent practice using repetitions of the same stimulus ("blocked" training) results in speedy acquisition, but lower levels of retention and generalization. Training using a set of related stimuli ("variable" training), while often producing slower acquisition, increases scores on retention and generalization tests. Schmidt and Bjork (1992) suggested that the evaluation of the overall success of training should be based on more than the speed at which the new skill was acquired. They believe that long term retention and transfer should also be considered as training criteria. Over the subsequent 16 years, 449 articles were published (56% of all articles on this topic since 1915), examining the benefits of variable training and testing theories attempting to explain its effectiveness.

One of the most common of these is Schmidt's (1975) schema theory. It states that practice using a single stimulus results in the encoding and storage of a specific stimulus-response unit, which makes retrieval of the

Author info: Correspondence should be sent to: the first author at 131 Krebs Hall, University of Pittsburgh at Johnstown, Johnstown, PA 15904; 814-269-2953; bertsch@pitt.edu
North American Journal of Psychology, 2009, Vol. 11, No. 3, 633-642.
© NAJP

response alone difficult when a novel stimulus is presented. Practice based on an array of different stimuli encourages the learner to abstract and store a general rule (schema) that describes the relationship between stimuli and behavior. This schema is created based on many types of information stored in memory during the learning process (some only temporarily), including environmental characteristics and the specific set of physical responses made. Once the schema is learned, it can be retrieved in response to any appropriate stimulus, including those not part of the initial training.

Other theories of procedural learning also propose that there are differences between the acquisition of stimulus specific knowledge, where one-to-one pairings produce automatic responses to stimuli (Logan, 1988; Schneider & Shiffrin, 1977) and strategic level knowledge, which (with practice) can be expressed independently from particular stimulus features (Anderson, 1982; Pellegrino, Doane, Fischer & Alderton, 1991).

Other types of generalization are also issues in training research. Studies on context generalization have indicated that an immediate shift to a new environment after training generally produces a decrease in the learned behavior. However, decades of animal studies have shown that context generalization is more likely over extended retention intervals (e.g., Perkins & Weyant, 1958). For example, Biedenkapp and Rudy (2007) found that footshock-induced fear training in rats produced context-dependent behavior after a single day, but that the contextual dependency was no longer evident after one week. This pattern (see Riccio, Ackil, & Burch-Vernon, 1992, for a review) has been traditionally explained as the result of context forgetting—the initial context that guided the acquisition of the new behavior is forgotten faster than the new behavior itself. This forgetting allows the behavior to be performed in contexts that were initially disruptive.

Bertsch and Lamb (2008) found evidence (among human participants) that instead of context forgetting per se, it was actually the relationship between a new behavior and the original contextual cues that appeared to weaken over time. They gave subjects practice reading reversed text under one set of context conditions (font size) and then tested them under another. Changing contextual features immediately after training slowed reading speed. After one week however, even though all participants were able to report that the font size had changed, their mean reading speed was no longer affected by that change. Thus, both the skill and the original context information were still represented in memory, but the ability of one to influence the other had faded.

We propose that stimulus generalization will follow the same pattern. We believe that during practice trials, a "falsework" is created between each stimulus experienced during training and the new behavior. Historically, a falsework is a temporary form used during construction to support new architecture. As the new structure becomes able to support itself, the falsework is removed, leaving the structure ready for use. We use the term similarly, to refer to the temporary binding of two types of information together during skill acquisition. These falseworks allow each stimulus to participate in guiding the acquisition of the skill, but are not required for its flexible, long term use. The strength (and therefore, the temporal stability) of any individual falsework is based on the number of times the stimulus and behavior are paired. We believe that repeated practice using a single stimulus (blocked practice) binds that stimulus tightly to the new skill. This strong falsework requires time to fade before generalization can take place. The same number of practice trials spread among several stimuli (variable practice) results in multiple weaker falseworks. These weaker connections allow for the flexible application of the new skill to novel stimuli immediately.

In the following experiments, we tested this theory using a mirror-tracing apparatus. Participants practiced tracing either a single geometric shape or a set of three different mirror-reflected shapes. They were then tested on a novel shape immediately, after two days, or after one week. The following hypotheses were tested:

> H_1: When tested immediately after training, the transfer performance of participants trained using a blocked schedule will be significantly slower compared to their performance on the final practice trial.
>
> H_2: When tested immediately after training, the transfer performance of participants trained using a variable schedule will not be significantly different from their performance on the final practice trial.
>
> H_3: After two days, the transfer performance of participants trained using a blocked schedule will be significantly faster compared to those tested immediately.
>
> H_4: After one week, the transfer performance of the participants trained using a blocked schedule will be significantly faster compared to those tested after two days.

METHOD

Participants

All 120 participants (Mean age = 19.31 years, *SD* = 4.10; 63% women) were student volunteers from the University of Pittsburgh at Johnstown, a small liberal arts college in western Pennsylvania. They were enrolled in Introduction to Psychology classes and received credit for participating in the study. All participants had normal or corrected-to-normal vision and had no previous experience with our research.

Materials

An apparatus that displayed the mirror-reflected image of a geometric figure (while preventing a direct view) was used to present the training and generalization test stimuli. Four different star shapes were used, three during training (4, 5 and 6 points) and one as the generalization test stimulus (7 points). Each was drawn in black marker and filled a white 4" by 6" card. A pencil sized stylus was used as the tracing tool and a handheld stopwatch was used to record the amount of time it took to trace around the shape. All participants began and ended at the same place on each shape.

Design & Procedure

A 2 (training schedule) x 3 (generalization test interval) between subjects factorial design was used. The training schedule was either Blocked, in which participants were trained using repetitions of the 5 pointed shape, or Variable, in which participants were trained using all three training stimuli, presented in a random order (with the exception that the same stimulus was not presented in consecutive trials). Generalization performance was tested using the 7 pointed shape immediately after training, after two days, or after one week. Data on two dependent variables was collected: the amount of time it took for a participant to trace around the shape, and the number of errors made.

Training. Each participant was randomly assigned to one of the two training schedules and was administered the training and testing individually. During the training phase, participants were asked to trace completely around the outline of the stimuli with the stylus, using only its reflection as a visual guide. They were asked to go as quickly as they could while making as few mistakes as possible. Each participant in the blocked schedule group received 10+1 consecutive practice trials (training was terminated after the first errorless trial following the tenth). The training criterion for those in the variable schedule group was increased by one trial (10+2), as in some cases the 4 pointed shape—which was easier to trace than the others—was randomly chosen on the 11th trial. Trials were observed by the researcher, who timed each using a hand-held stop watch, starting with the first movement of the stylus and stopping once the stimulus had been completely traced. Mistakes were also tallied, defined as failing to change the direction of the stylus correctly at the corners of the shapes. As these mistakes were self-evident to the participants no feedback was provided with regard to this variable. However, subjects were informed of their progress in tracing time.

Testing. The generalization testing was conducted in the same room, under the same conditions and with the same experimenter as the training phase. Each participant traced the 7 pointed shape twice and mean number of errors and mean tracing completion time was computed. In order to test our hypothesis that the stimulus used by those in the blocked training group remained present in memory even after it no longer influenced performance, we asked participants in the 1 week interval whether they noticed anything different about the test stimulus compared to the one on which they had practiced.

RESULTS

Due to our training criterion, significantly more training trials were required by the variable training group (M = 12.73, SD = 1.13) than by the blocked training group (M = 12.03, SD = 1.53), t (118) = 2.85; p = .005; η^2 = .06. An independent samples t-test indicated that the final training trial mean time was significantly faster for the variable trained group than for the block trained group, t (118) = 2.27; p = .02; η^2 = .04. This is likely to be due to the inclusion of the 4 pointed shape in the variable group's stimulus set—as it required the fewest course changes, it was completed faster. On average, participants in each training condition were approaching asymptotic performance in tracing time by the time they had reached their groups' (error-based) training criterion.

TABLE 1 Mean Tracing Time as a Function of Retention Interval and Training Condition.

	Training Condition			
	Blocked		Variable	
Retention Interval	M	n	M	n
Final Training Trial	23.23 (11.8)[a]	60	18.82 (9.4)[a]	60
Immediate	31.81[c] (1.85)[b]	20	23.47[c] (1.75)[b]	20
2 day	25.90[c] (1.71)[b]	20	25.02[c] (1.73)[b]	20
1 week	23.87[c] (1.75)[b]	20	27.45[c] (1.76)[b]	20

Note. [a] Standard Deviation, [b]Standard Error. [c] Mean adjusted for Final Training Trial time.

Compared to their mean time on the final training trial, participants in the blocked group were significantly slower at tracing the 7 pointed shape when tested immediately after training, t (78) = 4.82; p < .001; η^2 = .23. The immediate mean generalization time of the variable training group was not significantly different from their final training trial time.

A between subjects ANOVA was performed on the mean number of errors did not indicate significant differences between variable or blocked training groups, among the three retention intervals, or any interaction effects. In order to examine group differences in generalization test time, we first adjusted our group means by co-varying based on final practice trial time in order to equate the groups across this variable (Table 1). As the correlation between final practice trial time and generalization test time was .74 and the homogeneity of regression assumption was met, a between subjects ANCOVA was performed. While neither main effects of retention interval nor training group condition were significant, the interaction was, F (1, 113) = 5.60; p = .005, partial η^2 = .09 Further testing of the adjusted group means indicated a cross over type of interaction pattern, in that while those trained in the blocked condition significantly improved their generalization performance over time, F (2, 57) = 5.88, p = .005, those trained in the variable condition declined (although not significantly, F (2, 57) = 1.35). Among the blocked training groups, post hoc comparisons using Tukey's HSD indicated a significant difference between the immediate and 1 week retention intervals only. Finally, among the block trained participants tested after 1 week, 18 out of 20 reported that the test stimulus was different from the one on which they had trained.

DISCUSSION

We proposed that during training, temporary connections are created that bond the mental representation of the stimuli used to acquire a new skill to the procedural memory of the new skill itself. Our use of the term "falsework" to represent this bond is based on its use in architectural construction, where it refers to temporary supports created for building arches or molding concrete. When, over time, the architectural piece becomes self-sustaining, the falsework is then removed, leaving the permanent structure. We believe that a similar process can be found in the construction of new skills, where temporary features guide acquisition, but are not required for (and in fact, would impede) permanent use.

Our first two hypotheses were that in an immediate test of generalization, performance of participants trained using a blocked schedule would decline (relative to their final block of practice), while the performance of those in the variable condition would remain unchanged. Each of these was supported. When tested immediately, participants trained using repetitions of the same stimulus had significantly slower times when they were switched to a novel shape. Tracing speed of participants tested on a novel shape immediately after training on a set of stimuli was not significantly reduced. Our final two hypotheses refer to the changes in transfer test performance over longer retention intervals for participants trained using the blocked schedule. Individually, neither was supported, in that the incremental improvement from one retention interval to the next was not significant. However the larger picture is encouraging, as performance did steadily improve over time, with a significant overall difference between the immediate and 1 week intervals.

There is support in the literature for the idea that different components of memory are represented with different degrees of stability over time. In examining pattern recognition, Posner and Keele (1970) found that an abstracted prototype representing a collection of patterns was more accurately recalled over a 1 week retention period than were the original patterns themselves. In part based on this research, Schmidt's (1975) theory of motor learning also posited that, unlike the schema itself, the memory representations of the specific information used to construct the new motor schema were "…hypothesized to be relatively weak, so that rapid forgetting of them can be expected over time." (p. 240).

More recently, a transformational theory of memory storage has been proposed (Moscovitch et al., 2005; Winocur, Moscovitch & Sekeres, 2007), in which all initial learning is suggested to be context-dependent and therefore represented by hippocampal structures. Over time, some of that new content becomes divorced from the contextual cues present during learning and is thereafter stored in the neocortex as a more general type of schematic memory. Winocur et al. (2007) found support for this theory in that destruction of hippocampal structures in rats eliminated the context-dependent memory effects traditionally seen in immediate testing. The separation over time between target content and contextual information proposed by the transformational theory is congruent with the results obtained by Bertsch and Lamb (2008), who found that in human participants, context cues themselves remained stored in memory even when they were functionally separate from memory performance.

In addition to contextual cues, we believe that the stimulus information used to acquire a new skill also remains in memory and temporarily influences the efficiency with which that skill can be applied to new stimuli. In the current experiment 90% of the participants (in the blocked training condition) tested after 1 week reported that the test stimulus was different from that on which they had initially practiced. Despite this knowledge, the stimulus change no longer impacted performance. We believe that the increase in generalization accuracy over time found in the present experiment is the consequence of a change in the falsework relationship between training stimulus information and the new skill. These falseworks appear to be temporally unstable and as they weaken over time, generalization of the skill to novel stimuli is more likely. Even though in our study blocked training participants received significantly fewer practice trials, this schedule appeared to result in stronger falseworks—as evidenced by the longer tracing times needed for a novel shape immediately. Even when more trials were required to reach our training criterion, pairing several different stimuli with the skill during training appeared to create a set of weaker falseworks, which allowed effective generalization to take place immediately.

There are several avenues of inquiry that future research might explore. First, instead of using personal training criteria for each individual participant, a universal training criterion would likely show the offset of training time between blocked and variable training efficiency more clearly. Second, we did see a decline in transfer performance over time among those receiving variable training. Whether this was due to the amount of practice in this condition—and potentially therefore a reflection of falsework contributions to the initial learning process—should be tested. If variable practice creates weaker falseworks, perhaps they require more repetition to guide permanent acquisition.

Clearly there are many questions still unanswered in the quest for effective skill acquisition. Our research indicates that it may be difficult to "have it all"—the rapid training but delayed generalization ability afforded by a blocked training schedule, or the generally longer training period resulting in stronger immediate transfer performance enabled by variable training. If, as Schmidt and Bjork (1992) suggested, the definition of successful training is to be based in part on generalization efficiency, then different schedules of practice will need to be chosen not only with this goal in mind, but also with the time frame in which transfer performance would be required.

REFERENCES

Anderson, J. R. (1982). Acquisition of cognitive skill. *Psychological Review, 89,* 369-406.

Bertsch, S., & Lamb, A. (2008). Ensemble weakening and behavior generalization. *North American Journal of Psychology, 10,* 189-196.

Biedenkapp, J.C. & Rudy, J.W. (2007). Context pre-exposure prevents forgetting of a contextual fear memory: Implication for regional changes in brain activation patterns associated with recent and remote memory tests. *Learning & Memory, 14,* 200–203.

Logan, G. D. (1988). Toward an instance theory of automatization. *Psychological Review, 95,* 492-527.

Moscovitch, M., Rosenbaum, R. S., Gilboa, A., Addis, D. R., Westmacott, R., Grady, C., McAndrews, M. P., Levine, B., Black, S., Winocur, G., & Nadel, N. (2005). Functional neuroanatomy of remote episodic, semantic and spatial memory: A unified account based on multiple trace theory. *Journal of Anatomy, 207,* 35–66.

Pellegrino, J. W., Doane, S. M., Fischer, S. C., & Alderton, D. (1991). Stimulus complexity effects in visual comparisons: The effects of practice and learning context. *Journal of Experimental Psychology: Human Perception and Performance, 17*, 781-791.

Perkins, C. C. & Weyant, R. G. (1958). The interval between training and test trials as determiner of the slope of generalization gradients. *Journal of Comparative & Physiological Psychology, 51,* 596- 600.

Posner, M. I., & and Keele, S. W. (1970). Retention of abstract ideas. *Journal of Experimental Psychology, 83*, 304-308.

Riccio, D. C., Ackil, J. K., & Burch-Vernon, A. (1992). Forgetting of stimulus attributes: Methodological implications for assessing associative phenomena. *Psychological Bulletin, 112,* 433-445.

Schmidt, R. A. (1975). A schema theory of discrete motor skill learning. *Psychological Review, 82*, 225-260.

Schmidt, R. A. & Bjork, R. A. (1992). New conceptualizations of practice: Common principles in three paradigms suggest new concepts for training. *Psychological Science, 3,* 207-217.

Schneider, W. & Shiffrin, R. M. (1977). Controlled and automatic human information processing: I. Detection, search, and attention. *Psychological Review, 84,* 1-66.

Winocur, G., Moscovitch, M., & Sekered, M. (2007). Memory consolidation or transformation: Context manipulation and hippocampal representations of memory. *Nature Neuroscience, 10,* 555-557.

Copyright of North American Journal of Psychology is the property of North American Journal of Psychology and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.