Research Report

ASSOCIATION FOR PSYCHOLOGICAL SCIENCE

Psychological Science
2014, Vol. 25(9) 1816–1821

DOI: 10.1177/0956797614535401
pss.sagepub.com
SAGE

Stimulus Typicality Determines How Broadly Fear Is Generalized

Joseph E. Dunsmoor and Gregory L. Murphy
New York University

Abstract

The ability to represent knowledge at the category level promotes the transfer of learning. How this ability integrates with basic forms of conditioned learning is unknown but could explain why conditioned fear is overgeneralized after aversive experiences. We examined the impact of stimulus typicality—an important determinant of category-based induction—on fear learning and generalization. Typicality is known to affect the strength of categorical arguments; a premise involving typical exemplars (e.g., sparrow) is believed to apply to other members, whereas a premise about atypical exemplars (e.g., penguin) generalizes more narrowly to similar items. We adopted this framework to human fear conditioning and found that fear conditioned to typical exemplars generalized more readily to atypical members than vice versa, despite equal feature overlap across conditions. These findings have implications for understanding why some fearful events lead to broad overgeneralization of fear whereas others are regarded as isolated episodes.

Keywords
fear conditioning, category-based induction, generalization, skin conductance responses, reasoning

Received 1/29/14; Revision accepted 4/20/14

Classical fear conditioning is a basic form of associative learning found throughout the animal kingdom. In it, a *conditioned stimulus* (CS; e.g., a tone) paired with an aversive *unconditioned stimulus* (US; e.g., electrical shock) comes to elicit a defensive *conditioned response* (e.g., sweating). This process is essential to much behavioral control, from the fight-or-flight response to the avoidance of dangerous situations (Lang, Davis, & Öhman, 2000; LeDoux, 2000). It remains unclear, however, how this simple form of learning integrates with higher-order forms of reasoning in humans.

Humans induce unknown properties of an object on the basis of the known properties of another object. Knowledge of whole categories promotes such generalizations. For example, a fear of doctors may be generalized to nurses, hospitals, or medical equipment. Research on category-based induction investigates how people generalize properties using conceptual knowledge (Murphy, 2002). One relevant phenomenon of induction is typicality asymmetry (Osherson, Smith, Wilkie, Lopez, & Shafir, 1990). Consider the following arguments:

1. Mice have sesamoid bones. Therefore bats have sesamoid bones.
2. Bats have sesamoid bones. Therefore mice have sesamoid bones.

Although the arguments involve the same categories, Argument 1 is perceived as stronger than Argument 2, because typical items are better sources of generalization than atypical items are (Osherson et al., 1990; Rips, 1975). This asymmetry depends on the items' conceptual status, classifying them as mammals and evaluating their typicality.

Although conditioning is widely regarded as a simple form of learning governed by low-level systems, conceptual knowledge can be recruited when humans generalize fear. For example, Dunsmoor and his colleagues found that subjects who experienced shock paired with

Corresponding Author:
Joseph E. Dunsmoor, New York University–Psychology, 6 Washington Pl., New York, NY 10003
E-mail: joseph.dunsmoor@nyu.edu

Table 1. Animal Exemplars Used in the Experiment

Typicality	Birds	Mammals
Typical	Crow, hummingbird, sparrow	Bear, horse, rabbit
Atypical	Emu, ibis, penguin	Aardvark, armadillo, otter
Intermediate (unpaired controls)	Cockatiel, duck, dove, finch, macaw, owl	Bison, camel, deer, gorilla, panda, ram

some category members (e.g., various tools) generalized fear to other category members (Dunsmoor, Kragel, Martin, & LaBar, 2013; Dunsmoor, Martin, & LaBar, 2012). These results suggest that fear was associated to the concept "tool" as a whole rather than just the paired stimuli. However, an alternative explanation relies on feature overlap, as category exemplars tend to share perceptual features. The concept may have played no role.

To discover whether fear learning involves conceptual knowledge, we created a fear conditioning task that paralleled tests of typicality asymmetry in category-based induction. In a between-groups manipulation, subjects in two experimental groups underwent fear conditioning to pictures of either typical or atypical category members and then were tested for generalization of conditioned responses to pictures of either atypical or typical category members, respectively. For example, they might be conditioned to pictures of a crow, a hummingbird, and a sparrow (typical birds) and tested on pictures of an emu, an ibis, and a penguin (atypical birds)—or vice versa. Because the feature overlap of the learning and test items was identical in the two groups, any difference in generalization of conditioned responding would be due to conceptual properties rather than to overlap of perceptual features. Given the typicality asymmetry observed in studies of category-based induction, we hypothesized that fear conditioned to typical exemplars would generalize to atypical exemplars more readily than vice versa.

Method

Forty-six healthy volunteers provided written informed consent in accordance with the New York University Committee on Activities Involving Human Subjects and were compensated for their participation. Sample size was determined on the basis of a similar study in which category-based fear conditioning was measured by skin conductance response (SCR; Dunsmoor et al., 2012). Nine subjects were excluded because of an overall lack of measurable SCR, technical problems, or failure to follow instructions. The final sample included 37 subjects (26 women, 11 men; mean age = 23.4 years, *SD* = 4.8). Subjects were either conditioned to fear typical category members and then tested on atypical category members (typical-CS+ group: *n* = 19) or were conditioned to fear atypical category members and then tested on typical category members (atypical-CS+ group: *n* = 18). Subjects were assigned randomly to conditions.

All stimuli were black-and-white pictures of animals on a white background. They had been rated for typicality of category members by 22 raters on Amazon's Mechanical Turk, who used a scale from 1 (*highly atypical*) to 7 (*highly typical*). Mean ratings were 2.67 (*SD* = 0.68) for the atypical category members, 4.76 (*SD* = 0.41) for the intermediate category members, and 6.09 (*SD* = 0.37) for the typical category members. Each image appeared for 6 s; each presentation of an image was followed by a 10-s fixation cross.

During the conditioning phase, one animal category (e.g., birds), the CS+, was paired with a mildly aversive electrical shock to each subject's right wrist (US). The CS+ items included three highly typical or highly atypical basic level exemplars presented three times each (nine total CS+ trials), and each CS+ exemplar was paired with shock on two of three trials (66% reinforcement rate). An equal number of pictures from the other animal category (e.g., mammals), the CS−, were presented during fear conditioning and never paired with shock, which provided a within-subjects control. Unpaired category exemplars (CS−s) were of intermediate typicality. Table 1 lists the birds and mammals whose images were used as stimuli. Stimulus order and animal categories serving as CS+ and CS− were counterbalanced across subjects.

In a test of fear generalization immediately following conditioning, subjects viewed three novel exemplars from the CS+ category (e.g., new birds, referred to as GEN+, presented three times each) and three novel exemplars from the CS− category (e.g., new mammals, referred to as GEN−, presented three times each). No shocks were delivered during the generalization test; however, subjects were not informed that shocks would not occur and continued to make shock-expectancy ratings. For subjects conditioned to typical CS+ exemplars, GEN+ were atypical exemplars, whereas for subjects conditioned to atypical CS+ exemplars, GEN+ were typical exemplars. The GEN− exemplars were of intermediate typicality.

On each trial, subjects rated expectancy for a shock using a three-alternative forced-choice scale (1 = *no risk*, 2 = *moderate risk*, or 3 = *high risk*); this procedure was based on that described by Lissek et al. (2008). These ratings were used primarily to determine that subjects

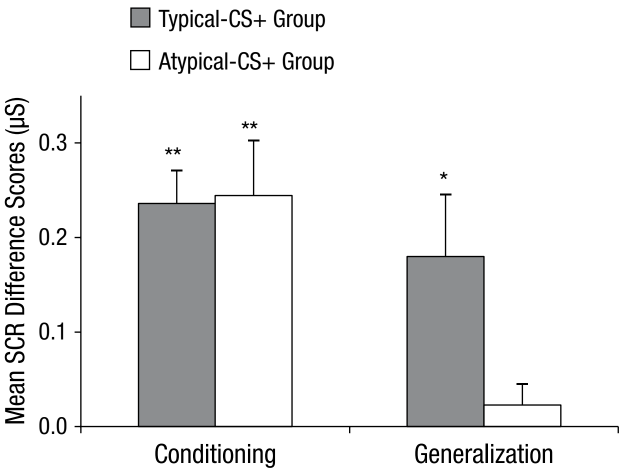

Fig. 1. Mean skin conductance response (SCR) difference scores for each experimental group and phase. For subjects in the typical-CS+ group, the conditioned stimuli paired with shock (CS+s) were images of typical animals (birds or mammals, counterbalanced), and the subjects were tested on new atypical animals from the same animal category (GEN+) in the generalization test. For subjects in the atypical-CS+ group, the CS+s were images of atypical animals, and the GEN+ stimuli were typical animals from the same category. See the Method section for an explanation of how the difference scores were calculated. Error bars represent 1 *SE*. Asterisks indicate scores significantly different from zero ($^*p < .05$, $^{**}p < .01$).

understood the CS-US contingencies, given that the CS+ exemplars varied throughout the conditioning phase, and CS-US pairing was intermittent. Shock ratings during conditioning were lost from 2 subjects because of technical problems with the presentation software in which button responses were not recorded; an additional subject failed to respond.

Shock was delivered using a stimulator from Grass Technologies (Warwick, RI), and the shock was calibrated separately for each subject to reach a level deemed highly annoying but not painful. SCR electrodes were placed on the hypothenar eminence of the palmar surface of the left hand and collected with a sampling rate of 200 Hz using a BIOPAC system (Goleta, CA). SCRs were scored according to criteria described previously (Dunsmoor et al., 2012; Green, Kragel, Fecteau, & LaBar, 2014) and square-root transformed before statistical analysis to normalize the distribution. The conditioned responses for the fear conditioning phase were calculated per subject as a difference score (CS+ minus CS–) using the mean SCRs to Trials 2 through 9 of the CS+ and CS– stimuli (Fig. 1). The first CS+ and CS– trials were excluded because learning had not yet occurred. The generalized conditioned response for the generalization phase was calculated per subject as the difference score GEN+ minus GEN– using the mean SCRs to Trials 1 through 9 of the GEN+ and GEN– stimuli.

Results

Both groups exhibited successful conditioning as revealed by greater SCRs to the CS+ than to the CS–: typical-CS+ group—$t(18) = 6.63$, $p < .001$; atypical-CS+ group—$t(17) = 4.12$, $p = .001$. There were no differences between groups in mean SCRs to the CS+ ($p = .26$) or CS− ($p = .21$). In short, learning was equal in the two groups. At the generalization test, the group trained on typical items generalized conditioned fear to atypical exemplars (i.e., GEN+ vs. GEN−), $t(18) = 2.79$, $p = .014$, but the group trained on atypical items showed no difference in response between GEN+ and GEN− items, $p = .34$. Mean SCRs to GEN+ items were also greater in the typical-CS+ group than in the atypical-CS+ group, $t(17) = 2.96$, $p = .009$; there were no group differences in mean SCRs to the GEN− items, $p = .17$. Figure 1 illustrates the mean SCR difference scores for each phase.

To examine the course of learning and generalization (see Fig. 2), we conducted analyses of variance using stimulus type and trial (1–9) as within-subjects factors and group (typical CS+, atypical CS+) as a between-subjects factor. During fear conditioning, SCRs showed a Stimulus Type (CS+, CS−) × Trial interaction, $F(8, 280) = 2.56$, $p = .01$, reflecting learning over trials, but no Stimulus Type × Group interaction, $p = .85$, and no Stimulus Type × Trial × Group interaction, $p = .30$. Shock-expectancy ratings during fear conditioning were similar across groups, as indicated by the fact that there was a Stimulus Type × Trial interaction, $F(8, 216) = 5.89$, $p < .001$, but no Stimulus Type × Group interaction, $p = .13$, or Stimulus Type × Trial × Group interaction, $p = .79$. These data demonstrate that learning, as measured by SCR and shock expectancy, was similar in the two groups.

Analysis of SCRs during the generalization test showed a Stimulus Type (GEN+, GEN−) × Group interaction, $F(1, 35) = 8.42$, $p = .006$, but no Stimulus Type × Trial interaction, $p = .98$, or Stimulus Type × Trial × Group interaction, $p = .48$. These results are in line with the effects shown in Figure 1, and group differences did not vary across trials. Shock-expectancy ratings during the generalization test were similar across groups; the Stimulus Type × Trial interaction was significant, $F(8, 256) = 4.05$, $p < .001$, but the Stimulus Type × Trial × Group interaction was not, $p = .13$, and the Stimulus Type × Group interaction only approached significance, $p = .053$.

Discussion

We have provided evidence that stimulus typicality influences generalization of conditioned fear in humans; fear was generalized from typical to atypical exemplars but was not generalized in the reverse direction, despite the

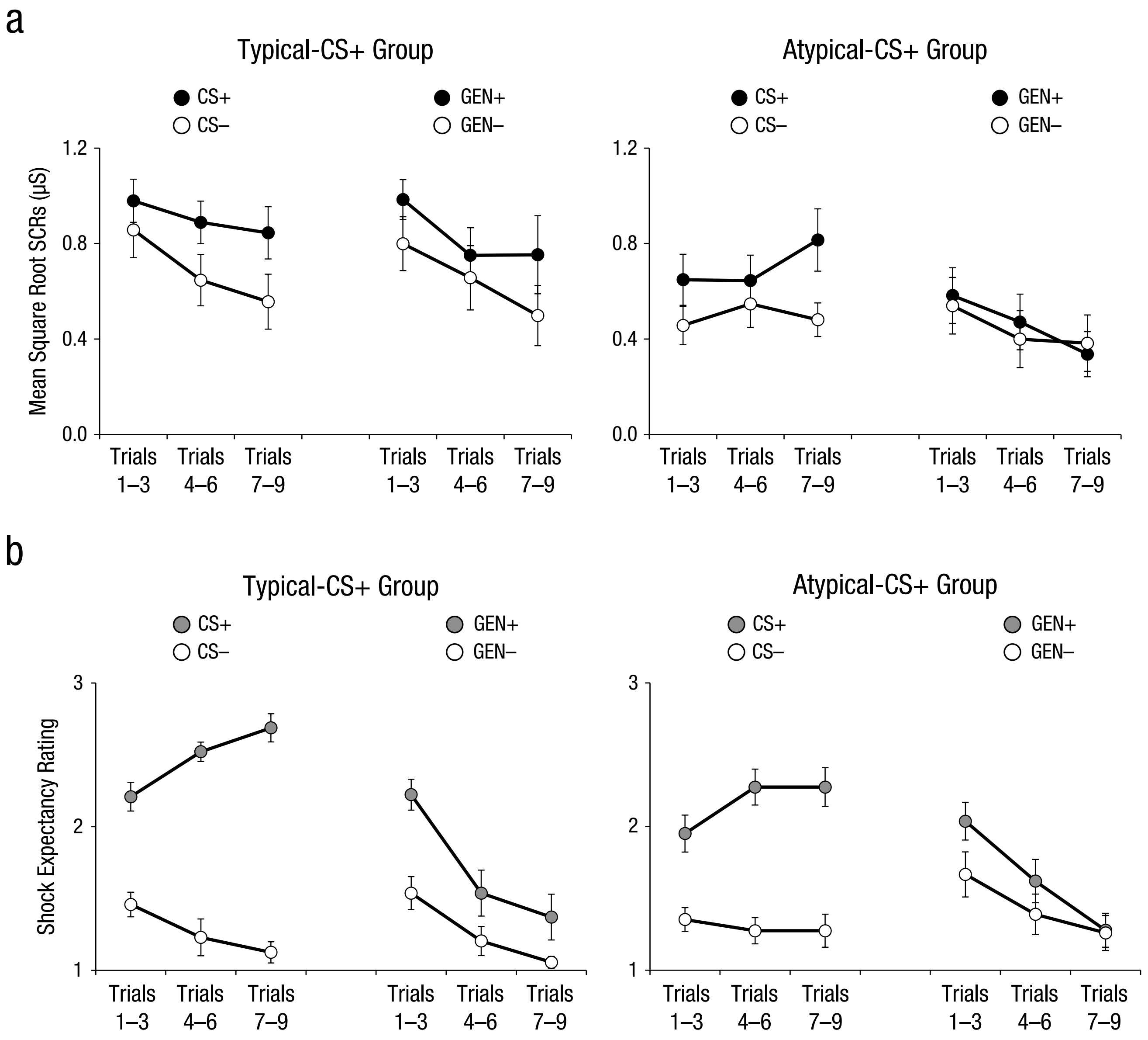

Fig. 2. Mean skin conductance response (SCR; a) and shock-expectancy rating (b) as a function of trial block, stimulus type, and experimental phase. Data are plotted separately for each group. Note that no shocks were delivered during the generalization test. Error bars represent ±1 *SE*. CS+ = animal category paired with shock; CS– = animal category never paired with shock; GEN+ = novel exemplars from the CS+ category; GEN– = novel exemplars from the CS– category.

fact that feature overlap did not differ between these two conditions. In the literature on category-based induction, this asymmetry is explained by the idea that typical items generate stronger expectations that the property holds for the entire category (Osherson et al., 1990). A similar mechanism could explain the fear generalization we observed—shock was more strongly associated to the concept (rather than to the specific learning exemplars) when typical items were used for conditioning than when atypical items were used. Because atypical items, such as penguins, were accepted as category members, conditioning to typical exemplars extended to atypical category members at test as a result of this stronger association. Although people can learn to associate atypical items with shock, this learning appears to be confined to the fear-conditioned exemplars, and fear expression does not as readily generalize from atypical items to other category members.

The primary question in this study concerned possible asymmetries in an autonomic measure of conditioned-fear expression (i.e., SCR). But the findings have implications for the potential contribution of higher-order cognitive processes to the generalization of conditioned fear. In this case, learning that typical CS+ exemplars predict shock may activate the internal structure of conceptual representations more than learning the same information about atypical CSs does. If a CS is a good representation of its category, pairing it with a shock may generate the proposition that other category members will be a potential threat as well (at least within the context of the fear-conditioning experiment). This explanation is in line with suggestions that associative learning in humans involves cognitive processes such as propositional and declarative knowledge about the CS-US relationship (Mitchell, De Houwer, & Lovibond, 2009). The present study adds novel evidence that typicality, a relatively complex attribute, influences generalization of fear learning and opens the way for further research on how other phenomena of category-based induction (conclusion specificity, premise diversity, mixed premises, etc.; Osherson et al., 1990) integrate with humans' acquisition and expression of fear.

The present experiment does not constitute a robust test of high-level reasoning processes in the course of learning and generalization. We provided subjects with a crude, three-alternative forced-choice scale for rating the perceived risk of shock on each trial and found no statistical evidence that these ratings were affected by stimulus typicality. Future research will be needed to investigate possible differences in fear generalization between autonomic and declarative measures.

These results also highlight different routes of stimulus generalization in humans that go beyond mere physical similarity. Fear-conditioning research almost universally employs simple sensory cues such as lights and tones. Many real-world fears are complex, however, and feared stimuli can often be represented by a network of interrelated concepts and information and can assume different forms from one encounter to the next. It remains a challenge to predict what information might acquire emotional significance and later trigger fear and anxiety in real-world situations. For example, combat veterans with posttraumatic stress disorder may show a categorical fear of certain objects reminiscent of the region in which they were deployed. Emerging research is beginning to reveal behavioral and neural mechanisms associated with fear generalization in humans (Dymond et al., 2011; Lissek et al., 2013; Vervliet, Kindt, Vansteenwegen, & Hermans, 2010). The present research suggests that how a person conceptualizes the CS may have a significant effect on how he or she generalizes fear to new stimuli. It therefore provides avenues for further study on how generalized fear learning is governed by semantic categories and their structure.

Author Contributions

J. E. Dunsmoor performed the testing and collected and analyzed the data. Both authors developed the study concept and design, interpreted the data, drafted the manuscript, and approved the final version of the manuscript for submission.

Acknowledgments

We thank Elizabeth Phelps for helpful comments.

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.

Funding

This research was supported by National Institute of Mental Health Training Award in Systems and Integrative Neuroscience Grant T32-MH019524 (to J. E. Dunsmoor) and by National Science Foundation Grant BCS 1128769 (to G. L. Murphy).

References

Dunsmoor, J. E., Kragel, P. A., Martin, A., & LaBar, K. S. (2013). Aversive learning modulates cortical representations of object categories. *Cerebral Cortex*. Advance online publication. doi:10.1093/cercor/bht138

Dunsmoor, J. E., Martin, A., & LaBar, K. S. (2012). Role of conceptual knowledge in learning and retention of conditioned fear. *Biological Psychology*, *89*, 300–305.

Dymond, S., Schlund, M. W., Roche, B., Whelan, R., Richards, J., & Davies, C. (2011). Inferred threat and safety: Symbolic generalization of human avoidance learning. *Behaviour Research and Therapy*, *49*, 614–621.

Green, S. R., Kragel, P. A., Fecteau, M. E., & LaBar, K. S. (2014). Development and validation of an unsupervised scoring system (Autonomate) for skin conductance response analysis. *International Journal of Psychophysiology*, *91*, 186–193.

Lang, P. J., Davis, M., & Öhman, A. (2000). Fear and anxiety: Animal models and human cognitive psychophysiology. *Journal of Affective Disorders*, *61*, 137–159.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual Review of Neuroscience*, *23*, 155–184.

Lissek, S., Biggs, A. L., Rabin, S. J., Cornwell, B. R., Alvarez, R. P., Pine, D. S., & Grillon, C. (2008). Generalization of conditioned fear-potentiated startle in humans: Experimental validation and clinical relevance. *Behaviour Research and Therapy*, *46*, 678–687.

Lissek, S., Bradford, D. E., Alvarez, R. P., Burton, P., Espensen-Sturges, T., Reynolds, R. C., & Grillon, C. (2013). Neural substrates of classically conditioned fear-generalization in humans: A parametric fMRI study. *Social Cognitive and Affective Neuroscience*. Advance online publication. doi:10.1093/scan/nst096

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning [Target article and commentaries]. *Behavioral & Brain Sciences*, *32*, 183–246.

Murphy, G. L. (2002). *The big book of concepts*. Cambridge, MA: MIT Press.

Osherson, D. N., Smith, E. E., Wilkie, O., Lopez, A., & Shafir, E. (1990). Category-based induction. *Psychological Review*, *97*, 185–200.

Rips, L. J. (1975). Inductive judgments about natural categories. *Journal of Verbal Learning and Verbal Behavior*, *14*, 665–681.

Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010). Fear generalization in humans: Impact of prior nonfearful experiences. *Behaviour Research and Therapy*, *48*, 1078–1084.