UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Effects of Discrimination Difficulty on Peak Shift and Generalization

Permalink
https://escholarship.org/uc/item/4br7d85x

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 46(0)

Authors
Lee, Jessica C.
Cahyadi, Tamara
Lovibond, Peter
et al.

Publication Date
2024

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Effects of Discrimination Difficulty on Peak Shift and Generalization

Jessica C. Lee (jessica.c.lee@sydney.edu.au)
University of Sydney; University of New South Wales Sydney

Tamara Cahyadi (tamaralarissac@gmail.com)
University of New South Wales, Sydney

Peter F. Lovibond (p.lovibond@unsw.edu.au)
University of New South Wales, Sydney

René Schlegelmilch (r.schlegelmilch@uni-bremen.de)
University of Bremen

Abstract

In this paper, we test the effect of manipulating discrimination difficulty on subsequent generalization of learning and in particular, on the peak shift effect. Participants learned a discrimination where one stimulus led to an outcome (S+) and another stimulus led to no outcome (S-). Difficulty was manipulated by varying the degree of similarity between the S+ and S- across groups (easy/medium/hard). In contrast to similar studies in animals, we found that increasing the difficulty of the discrimination resulted in less peak shift. Using a hierarchical mixture model, we characterize the effects of discrimination difficulty on relational- and similarity based responding, and show for the first time, a similar mixture of responding on stimulus identification gradients. We conclude that peak shift on generalization and identification measures can be explained by mixtures of participants responding in different ways.

Keywords: generalization; discrimination; peak shift; identification; similarity

Introduction

Stimulus generalization refers to the adaptive ability of animals and humans to transfer learned responses from familiar stimuli to novel stimuli. In associative learning, the question of how this is achieved has been the topic of theoretical and empirical investigation for over 50 years. In everyday life, we often have to integrate learning about multiple stimuli in order to make accurate predictions. We can learn about stimuli that predict an outcome (S+) as well as stimuli that do not predict that outcome (S-). As the physical difference between the S+ and S- becomes smaller and smaller, discriminating between the stimuli to make accurate predictions becomes more difficult. In this paper, we address the question of how the difficulty of the discrimination (or the similarity between S+ and S-) affects subsequent generalization to novel stimuli.

The first study addressing this question was by Hanson (1959), who trained groups of pigeons by reinforcing a pecking response at a keylight of a specific color (S+). One group of pigeons were exposed to the S+ only, while other groups of pigeons received additional exposure to a keylight where responding was not reinforced (S-), at varying degrees of similarity to the reinforced S+ keylight. In other words, the discrimination difficulty (physical similarity between the S+ and S-) differed between groups. To assess generalization, the color of the keylight was varied along the wavelength dimension and the amount of responding was measured. The group that only received exposure to the S+ showed a peaked, symmetrical gradient with the highest responding at the S+ value. The groups that had received discrimination training with an additional S- showed a *peak shift* (for a review, see Purtle, 1973), where the peak of the gradient shifted from the S+ to a value further along the dimension away from the S-. The degree to which the peak shifted along the dimension was proportional to the difficulty of the discrimination the pigeons had learned. The harder the discrimination (i.e., the more similar the S+ and S-), the larger the peak shift.

Peak shift has broad interest beyond conditioning procedures as it describes a phenomenon that can occur whenever there are two known stimuli with different outcomes producing a maximal response for a different stimulus with more exaggerated features. Peak shift has been used in ethology to explain sexual imprinting (ten Cate et al., 2006) and signal evolution (Lynn et al., 2005). Such shifts in responding can be seen as adaptive since they minimize the likelihood of misidentifying important stimuli such as predators or potential mates. Humans also display effects reminiscent of peak shift in various perceptual and cognitive domains. For instance, caricatures of faces are easier to recognize than actual faces (e.g., Lewis, 1999), shifts in category representation ("idealization" effects) occur as a result of contrast (Davis & Love, 2010), and we reproduce phonemes in a way that exaggerates their differences compared to casual speech (the hyperspace effect, Johnson et al., 1993). The pervasiveness of peak shift across a range of tasks and species suggests that the same underlying learning and generalization mechanisms underlie basic forms of discrimination learning as well as the formation of more abstract, conceptual knowledge.

In comparison to the extensive literature documenting the peak shift phenomenon, there has been relatively little research examining the relationship between discrimination difficulty and peak shift. Hanson's (1959) original report of larger peak shifts with harder discriminations has been replicated in animals by Thomas (1962), who additionally found steeper generalization gradients with harder discriminations. Furthermore, the effect (more peak shift with increasing discrimination difficulty) is predicted by associative models that assume that generalization results from the interaction of ex-

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.* ©2024 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

citation from the S+ and inhibition from the S- (e.g., Blough, 1975; Ghirlanda & Enquist, 1998; McLaren & Mackintosh, 2002). The effect has also been found in humans (e.g., Baron, 1973; Derenne, 2006), however there have also been some failures to replicate (e.g., Doll & Thomas, 1967), and some evidence supporting the opposite relationship, with greater peak shift with easier discriminations (e.g., Thomas et al., 1973, 1991). Thomas (1993) has explained the contradictory results by proposing that humans use a relational response strategy involving comparison of the current stimulus with the adaptation level, which is the average of all presented stimuli. Participants then respond in different ways depending on the relation (e.g., respond if bluer than average, do not respond if greener than average).

J. C. Lee et al. (2018) also showed that relational learning was critical in discrimination learning and in particular, for obtaining peak shift. They found that one subgroup of participants generalized on the basis of physical similarity to the S+, while another subgroup of participants generalized on the basis of relations between the stimuli (e.g., "the bluer the stimulus, the higher the likelihood of the outcome"), and these gradients averaged to form a peak-shifted gradient at the aggregate while each subgroup showed markedly different gradient shapes. Therefore, peak shift can be accounted for by mixtures of similarity and relational rules (see also J. Lee et al., 2021). One implication of this account is that the effect of discrimination difficulty on peak shift should be mediated through its effect on similarity and relational rules. To test this prediction, we conducted an experiment and used a mixture model to separately estimate similarity- and relational-based responding.

Experiment

The primary aims of the study were: 1) to test the effect of manipulating discrimination difficulty on generalization gradients and specifically, the magnitude of peak shift, and 2) to test whether and how discrimination difficulty affects the mixture of similarity and relational rules in generalization gradients. The evidence regarding the effect of discrimination difficulty in humans is mixed, and is complicated by the fact that the tasks that are typically employed are distinct from a conditioning task where predictive associations are learned between stimuli and outcomes. For example, Baron (1973) instructed participants to make a response whenever they heard the "correct" tone (i.e., the S+), which is more like an identification response than a conditioned response. For this reason, we also included an identification test following our generalization test, similar to that used by Lovibond et al. (2020). Thus, our third aim was to test whether similar mixtures of responding could be found in identification gradients, and thus, whether peak shift in identification was subject to the same explanation proposed by J. C. Lee et al. (2018).

Method

Participants Two hundred and twenty-two Psychology students (M age = 19.8, SD = 3.5, 143 females, 78 males, 1 other) participated in exchange for partial course credit. The total number of participants tested in each group were 68 in the easy group, 62 in the medium group, and 92 in the hard group. Additional participants were recruited for the hard group since initial testing indicated that a larger proportion of participants were failing the training criterion.

Materials The experiment was programmed using jspsych (De Leeuw, 2015), hosted using JATOS (Lange et al., 2015), and run online with participants' own computer, mouse, and keyboard. The stimuli were the same as Lovibond et al. (2020) comprising 21 colored rectangles varying between green (hue value 0.4) to blue (hue value 0.6) in equal increments, with saturation and brightness fixed at 1 and .75 respectively.

Procedure The experiment was similar to Lovibond et al. (2020), with an initial training phase, generalization test, rule assessment, and an identification test. Participants were randomly assigned to group Easy, Medium, or Hard. We manipulated discrimination difficulty by fixing the identity of the S+ along the dimension (at the midpoint) and including different S- in each group. Depending on group, the S- was either 2 (Hard), 4 (Medium), or 6 (Easy) steps away from the S+ (see Figure 1). The direction of the dimension, and therefore whether the S- was greener or bluer than the S+, was counterbalanced between participants. A predictive learning scenario was used where participants were asked to make predictions about whether a hypothetical shock outcome would occur based on a "signal" (i.e., the stimuli) that would appear on the shock machine.

On each trial during the training phase, a stimulus (either the S+ or S-) was presented, and 500ms later, text appeared asking "What do you think will happen? SHOCK - press L or NO SHOCK - press A". After participants responded by pressing A or L (no other keys would register), the stimulus remained on screen and the text was replaced with feedback, either "No Shock" text or "SHOCK!" text accompanied by a picture of a lightning bolt. Feedback was presented for 2 seconds and each trial in this, and all subsequent phases, was separated by a 2 second blank inter-trial interval (ITI). The S+ and S- were each presented 12 times each, with the S+ followed by the shock 75% of the time, and the S- never followed by the shock. Trials were randomized in blocks of 4 presentations of each stimulus, with the first S+ presentation of each block always followed by the shock.

Participants then proceeded to the generalization test, where they were asked to continue making predictions about the shock outcome, but with feedback withheld. Participants were presented with all 21 stimuli on the blue-green dimension once in randomized order, and made their ratings by clicking on a visual analogue scale ranging from "Certain NO SHOCK (0% chance of shock)" to "Certain SHOCK (100% chance of shock)", and then clicking on "Continue" to proceed to the next trial.

After the generalization test, participants were given a three-alternative forced-choice question. Participants could select from a similarity rule ”The more SIMILAR the color to AQUA (half blue/half green), the higher the likelihood of SHOCK”), and two relational rules (”The GREENER/BLUER the symbol, the higher the likelihood of SHOCK”).

Participants then completed the identification test, where all 21 stimuli were again presented once in randomized order. On each trial, participants were asked ”Is this symbol: the SAME as the one that led to SHOCK in Phase 1 (press S) or is it DIFFERENT (press D)”. Finally, participants were asked to indicate whether they suffered from any form of color-blindness by clicking ”yes” or ”no”.

Results

The data, code, and supplemental figures are available at osf.io/6f25s/.

Exclusion Criteria Participants were excluded if they indicated that they were color-blind, if they selected the inconsistent relational rule in the forced-choice question, or if they failed to pass the training criterion (average accuracy for the CS+ and the CS- had to be $> 50\%$). After exclusions, there were 53 participants in the easy group, 52 in the medium group, and 52 in the hard group (N=157 total).

Acquisition For brevity, the acquisition data will not be presented. All three groups showed clear discrimination in their predictions for the stimuli, with some differences in speed of acquisition and terminal performance (expected given the manipulation of discrimination difficulty).

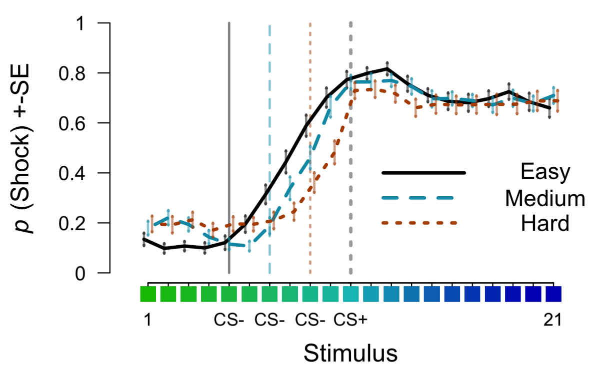

Figure 1: Aggregate generalization gradients in each group.

Modelling

Figure 1 shows the aggregate generalization gradients. In contrast to the animal data, there is more apparent peak shift with easier discriminations. To analyze the data, we employ a latent class hierarchical modeling approach (see M. D. Lee & Wagenmakers, 2014) used in previous research (J. C. Lee et al., 2021; Schlegelmilch et al., 2023). As Figure 2 illustrates, we assume a mixture of populations (rule vs. similarity), and a model selection procedure estimates for each individual k, whether their gradient is purely rule-like (Sigmoid/standard logistic function), or whether it should be augmented (i.e., combined) with a Gaussian downward trend once it surpasses a specific color near the CS+ (i.e., if shock is less likely the more distant from CS+). The selected function's prediction is then applied to the data via a Beta regression.

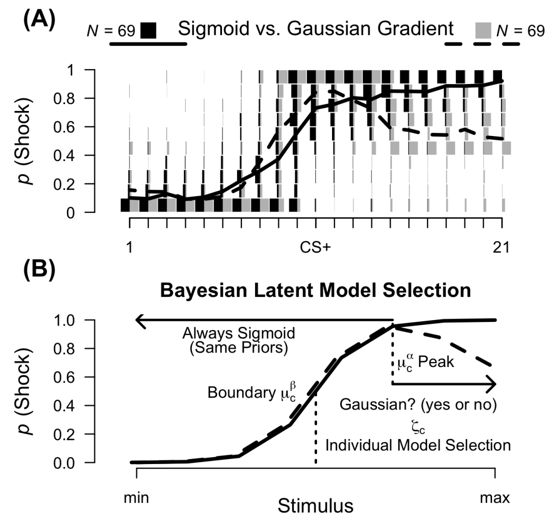

Figure 2: A) Bayesian model classification procedure and B) generalization gradients in Sigmoid and Gaussian subgroups (histograms = individual responses).

In this model (Fig.2B), we estimate three key dependent population variables. The first variable is the probability of strategy assignment $\zeta_c \sim \text{Gaussian}(0,1)$ (log-scale) in a given condition c, and a categorical model index is sampled for each individual, $z_{kc} \sim \text{Binomial}((1+\exp(\zeta_c)^{-1}))$ (0=rule-like vs. 1=similarity). If discrimination difficulty qualitatively affects how participants render the shock estimates, this would change ζ_c.

The second variable is the boundary between CS+ and CS$-$ defining the location at which participants respond with 50% shock. This is captured by the parameter $\mu_c^\beta \sim$ Gaussian(0,1), sampled hierarchically at the individual level[1]. Specifically, this parameter is part of the Sigmoid function which we apply to all participants. However, if an individual's model assignment is $z_{kc} = 1$, the Sigmoid predictions will be replaced by the Gaussian downward trend beginning at the estimated peak location $\mu_c^\alpha \sim$ Gaussian(0,1), which is the third variable of interest. This means that the boundary μ_c^β is estimated based on all participants, but μ_c^α is only estimated for those for whom the Sigmoid is augmented by the Gaussian downward trend.

Finally, we applied the probability estimates $\hat{p}_{ks}$ for each

[1]The stimulus values are centered on CS+ $= 0$ in the model, with standardized distances of 0.161 between stimuli (min/max = $-+1.61$). Negative estimates = boundary left-hand to CS+.

stimulus s and participant k, to the data y_{ks} via Beta regression, $y_{ks} \sim \text{Beta}(4 \cdot \hat{p}_{ks}, 4 \cdot (1 - \hat{p}_{ks})$. Note that the technical implementation is slightly more complex than illustrated, but we omit these details for brevity (see osf.io/6f25s/ for further details). We implemented the model using R (R Core Team, 2018) and the package JAGS (Plummer, 2023). It converged on four chains in 20000 iterations with sampler adaptation in 5000 samples.

For testing between-group differences, we freely estimated the Bayesian hyper posteriors ζ_c, μ_c^β, and μ_c^α for each. We then subtracted the Bayesian samples pair-wise between groups to obtain mean differences and their 95% highest-density intervals (HDIs). If an HDI excludes 0, we interpret this as evidence for an effect.

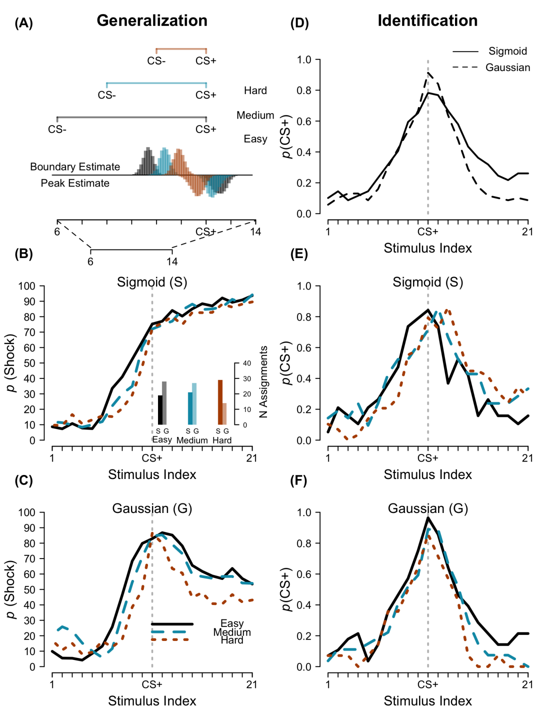

Figure 3: Generalization and Identification Results after Bayesian Model Assignment. See text for explanation.

Generalization Phase As can be seen in Figure 2A, 69 participants were classified into each strategy. 19 participants were classified as guessing and were omitted from the analyses. The individual assignments exceeded 80% certainty (number of consistent z_{kc} samples) for 93% of the participants. Of those participants verbally reporting a relational rule strategy 67.5% were assigned to the Sigmoid, and of those who reported a similarity strategy 74.1% were assigned to the Gaussian, underlining the credibility of the modeling procedure. Figure 3 illustrates the expectancy gradients for each group, and assigned model, together with the group-level posteriors.

The bar graphs in Figure 3B show that the number of rule participants (Sigmoid) increased with difficulty. The assignment probability (ζ_c converted to probability p) increased from the easy (E: $p = .415$) to the hard condition (H: $p = .655$), $M_{E-H} = -0.24$ 95%HDI[-0.43,-.04]. While the probability fell in-between both in the medium condition (M: $p = .456$), the evidence for its difference to the easy and hard condition was inconclusive, with $M_{E-M} = -0.041$ 95%HDI[-0.24,.16] and $M_{M-H} = -0.20$ 95%HDI[-0.40,.01], respectively. Note, if increased difficulty had led to a stronger shift of the peak towards the end-point of the dimension, this could increase the likelihood of Sigmoid classifications, because near the end-point a Gaussian downward trend is impossible. However, this would result in differences in the Sigmoid gradients between the three conditions (i.e., more gradual/less steep slopes in the hard compared to the other conditions), which was not the case (see Figure 3B).

Figure 3A shows the Bayesian posteriors for the boundary estimates. As can be seen, they are orderly and correspond to the difficulty manipulation, with $M_E = -.38$, $M_M = -.27$, and $M_H = -.18$. There was evidence for a difference in all three comparisons, $M_{E-M} = -0.109$ 95%HDI[-0.20,-.02], $M_{E-H} = -0.201$ 95%HDI[-0.29,-.11], and $M_{M-H} = -0.092$ 95%HDI[-0.19,0]. These estimates reflect boundary shifts aggregated over strategy (Sigmoid and Gaussian). Thus, regardless of strategy, moving the CS- away from the CS+ resulted in proportional shifts of the gradient boundary.

Regarding the peak estimates in the Gaussian group, there was an orderly shift between $M_E = .084$, $M_M = .058$, to $M_H = -.045$, with evidence for a difference between easy and hard conditions, $M_{E-H} = -0.143$ 95%HDI[-0.26,-.03]. Together with the previous result, this means that shock-expectancy narrowed around the CS+ in the hard, compared to the easy group (Figure 3C), which can also be seen in the average Gaussian gradients (Figure 1). The evidence for the other comparisons was inconclusive, with $M_{E-M} = -0.039$ 95%HDI[-0.15,.08], and $M_{M-H} = -0.104$ 95%HDI[-0.22,.01].

Identification Phase Figure 3D shows the identification gradients plotted by the previously assigned generalization model (Sigmoid vs. Gaussian). Participants in the Sigmoid cluster tended to show an area shift (more 'same' responding beyond the CS+), suggestive of a carry-over effect from the previous generalization test (see Lovibond et al., 2020), while the Gaussian cluster did not. Figures 3E & F, further show that there was a tendency of a peak shift in the Sigmoid group from the easy to medium and hard conditions, which was not the case for the Gaussian group.

However, these average gradients hide the fact that the individual patterns were again the result of two distinct subgroups responding in different ways. Specifically, we found that some participants responded with 'different' (0) for stimuli on the left of the CS+ but with 'same' (1) for those on the right of the CS+ in a *Step* function, while other participants responded with 'same' for about 2 to 6 stimuli near the CS+ otherwise 'different', which we called *Plateau* responding. To analyze these subgroups, we used a similar but non-hierarchical cluster analysis as before (there was only one 0 or 1 response per stimulus and participant). Instead of a Sigmoid, we defined a Step function estimating the stimulus location $l^{S} \sim \text{Gaussian}_{[-1,1]}(0,1)$ below which 'same' responding was $p = .05$ and above which it was $p = .95$. Instead of augmenting a Gaussian, we defined a Plateau function estimating two locations $l_1^{P} \sim \text{Gaussian}_{[-1,.1]}(0,1)$ and $l_2^{P} \sim \text{Gaussian}_{[-.1,1]}(0,1)$, such that responding in-between was set to $p = .95$, and outside to $p = .05$. Otherwise we used the same classification method as in the Generalization analysis, and estimated the three parameters for each difficulty group by previously assigned model (Sigmoid vs. Gaussian) separately, which converged on 5000 iterations on four chains.

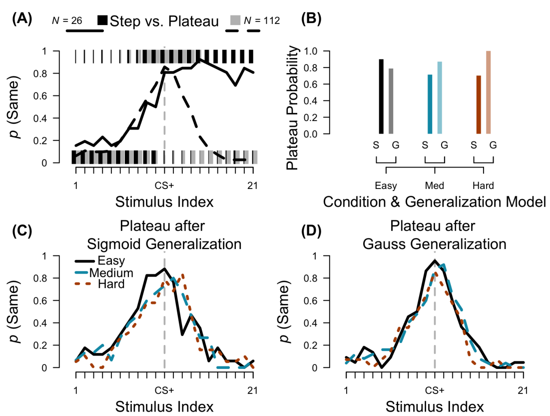

Figure 4: Identification Patterns by Generalization Model and Condition. See text for explanation.

Figure 4A shows that 26 participants were assigned to Step, and 112 to the Plateau function. Although in the minority, it appears that the Step cluster is responsible for the area shift seen in Figure 3D, which can be seen as a carry-over effect from the generalization test. Indeed, 18 of the 26 Step participants were previously assigned to the Sigmoid function (69%), supporting this suspicion. Figure 4B shows that the probability of Plateau responding was highest (100%) when assigned to the Gaussian in the hard condition, and this probability was lower in both medium and easy condition. Interestingly, the reverse seemed to be the case when the previously assigned model was the Sigmoid function. Still the evidence for these effects was generally inconclusive in pairwise comparisons.

Focusing on the Plateau group, we found an interaction between previous Generalization assignment and group regarding the estimated Plateau bounds l_1^{P} (left-hand) and l_2^{P} (right-hand). Only when participants were previously assigned to the Sigmoid, did the l_2^{P} estimates in the easy $M_E = .24$, medium $M_M = .41$, and hard groups $M_H = .40$, differ between easy and medium, $M_{E-M} = -0.166$ 95%HDI[-0.30,-0.04], and easy and hard, $M_{E-H} = -0.163$ 95%HDI[-0.29,-0.04]. This means that peak shift in CS+ identification depended on the Generalization pattern (Sigmoid or Gaussian), seen in Figure 4C.

In addition, the l_1^{P}in the Sigmoid cluster, easy $M_E = -.56$, medium $M_M = -.44$, and hard groups $M_H = -.40$, showed an additional area shift from the easy to hard condition, $M_{E-H} = -0.160$ 95%HDI[-0.285,-0.03], while all other comparisons included zero in their HDI's. Thus, there is evidence that if participants responded in a rule-like fashion during generalization they also showed a peak shift in the identification task, (a) in terms of a carry-over effect (Step function), but also (b) when they later showed Plateau responding in terms of an actual peak-shift. However, if participants in generalization showed a Gaussian similarity-like gradient, there was hardly any evidence of peak shift and CS+ identification was quite accurate (Figure 4D).

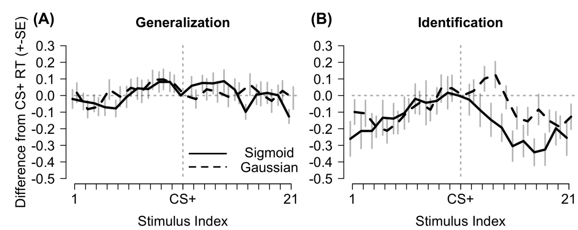

Figure 5: Response Times (RT) in Generalization and Identification. Y-axis represents RT(Stimulus) minus RT (CS+) calculated on the log-scale (positive = slower than for CS+). Error bars = +-1SE. Gray vertical line = CS+ location, gray horizontal = zero difference reference point.

To gain deeper insight into the difference between Gaussian and Sigmoid subgroups, we conducted an exploratory analysis on the reaction times in both tests. Based on the simple idea that identification RTs should become slower for stimuli more similar to the CS+, we calculated a difference score, subtracting the RT for the CS+ for each stimulus per individual. Figure 5 shows these difference scores aggregated across subgroups. RT's for identification (Figure 5B) were generally faster with increasing distance to the CS+ for the Sigmoid subgroup. However, the Gaussian generalization group responded more slowly to items close to the CS+ on the right side, i.e., in the region where one would usually expect a peak shift. Intriguingly, no such interactions were present for RTs in the generalization test (Figure 5A). Theoretically, this might suggest that the Gaussian subgroup needed extra effort to *suppress* 'same' responding for stimuli beyond the CS+ in order to avoid showing a peak shift in the identification task. Further studies are needed to confirm our speculations.

Discussion

In this study we tested the effect of discrimination difficulty (easy vs. medium vs. hard) on stimulus generalization and identification by manipulating the similarity of the S+ and S- along a blue-green color dimension. Harder discriminations are known to generate more shift in the peak of the generalization gradient in animals, but results are more mixed in humans. One explanation for this discrepancy is that previous studies have shown that humans show mixtures of similarity- and relational rule learning (J. C. Lee et al., 2018), and thus we also tested the effect of discrimination difficulty on rules using a mixture model. Our classification of participants into Sigmoid or Gaussian subgroups showed good correspondence with verbal report of Similarity and Relational rules used in previous studies (J. C. Lee et al., 2018).

At the aggregate level, we found that: 1) in contrast to studies on animals, the gradients showed more peak shift with easier discriminations, and 2) more difficult discriminations resulted in more participants classified as Sigmoid (i.e., using relational rules). These effects were most apparent when comparing groups Easy and Hard. Increased relational rule use in the Hard condition makes sense if we assume that more attention to the relational features of the stimuli is needed in order to discriminate between the stimuli. More attention to relational features during training might then mean that participants generalize on the basis of these features at test. If we assume that relational and physical features compete for attention, this might also explain why identification performance was poorer for the S+ in the hard group.

Turning to the effects of discrimination difficulty on the generalization subgroups (Sigmoid vs. Gaussian), we found that: 1) easier discriminations pulled the boundary of the Sigmoid toward the CS- but did not affect the slope, and 2) easier discriminations resulted in the peak of the Gaussian being further away from the CS- but did not affect the other Gaussian parameters. These results support the observation of greater peak shift with easier discriminations apparent in the aggregate gradients, but allows us to isolate this finding to the Gaussian subgroup (who generalized according to similarity).

Interestingly, consistent with the pigeon data reported by Thomas (1962), we also found narrower generalization gradients with harder discriminations, but only for participants classified as Gaussian. Another interesting observation is that for the Sigmoid participants, there was little evidence of group differences beyond the CS+ and CS-. In other words, manipulating discrimination difficulty seems to affect interpolation (responding between the known CS+ and CS-) but not extrapolation (responding beyond known stimuli).

The fact that the aggregate results were driven by the Gaussian subgroup has important theoretical implications. If the Gaussian subgroup had shown results consistent with animals, this would suggest that humans are capable of learning and generalizing associatively, but this behavior is obscured by sigmoidal gradients exhibited by other participants who learn about rules/relations. Instead, we found the opposite result to animals, with larger peak shifts with easier discriminations in the Gaussian subgroup. At present, we can only speculate about the reason for this result. The relational response strategy (Thomas, 1993) discussed earlier could explain the peak shift results since there is a larger shift in adaptation level over the course of testing (S8 to S11) for easier discriminations. This account does not however, explain why the gradients gets sharper as the discrimination becomes harder. Differences in perception are highly unlikely since there were no corresponding group differences in identification gradients. Further studies are needed to understand why both animals and humans show sharper gradients after hard discriminations, but the opposite pattern with regards to the relationship between peak shift and discrimination difficulty.

A novel result in our study was that we found a similar mixture of response patterns (Step vs. Plateau) in the identification test, with partial overlap between the subgroups classified from the generalization test (Sigmoid vs. Gaussian). It appears that carry-over effects from generalization to identification tests reported by Lovibond et al. (2020) are primarily driven by relational rule participants (Sigmoid) who continue to respond similarly in the identification test. This finding has implications for interpreting human studies purporting to demonstrate peak shift using an identification task where participants are instructed to make a response to a target stimulus. If a minority of participants misinterpret their task or inappropriately apply a relational rule, then the overall gradient may become peak-shifted. Our results suggest that when participants do the identification task properly (i.e., if we exclude Step participants), there is no area or peak shift in identification gradients. Thus, any peak shift displayed in an identification task may be explainable by individual differences in rules or response strategy, the same explanation proposed by J. C. Lee et al. (2018) to explain peak shift in generalization.

A potential limitation of the study was that a much larger proportion of participants were excluded in the Hard group compared to the Easy and Medium groups. While the training criterion was necessary to ensure that we excluded non-learners from all three groups, it may have resulted in a selection bias in the final sample, whereby the Hard group consisted of more attentive learners. However, the shape of the gradients did not change when we examined the full sample.

In conclusion, peak shift in humans does not conform to the predictions of associative models regarding the effect of discrimination difficulty. In contrast to animals, humans tend to display more peak shift with easier discriminations. We found that discrimination difficulty affects generalization strategy (more rule responders), as well as peak location for those classified as similarity responders (less peak shift). We have also shown for the first time, that this mixture of response strategies also applies to identification gradients. Thus, the evidence for peak shift in humans as conceptualized by associative models, may be more elusive than originally thought.

References

Baron, A. (1973). Postdiscrimination gradients of human subjects on a tone continuum. *Journal of Experimental Psychology*, *101*(2), 337.

Blough, D. S. (1975). Steady state data and a quantitative model of operant generalization and discrimination. *Journal of Experimental Psychology: Animal Behavior Processes*, *1*(1), 3.

Davis, T., & Love, B. C. (2010). Memory for category information is idealized through contrast with competing options. *Psychological Science*, *21*(2), 234–242.

De Leeuw, J. R. (2015). jspsych: A Javascript library for creating behavioral experiments in a Web browser. *Behavior Research Methods*, *47*(1), 1–12. doi: 10.3758/s13428-014-0458-y

Derenne, A. (2006). Effects of s+ and s-separation on gradient shifts in humans. *The Journal of general psychology*, *133*(2), 163–173.

Doll, T. J., & Thomas, D. R. (1967). Effects of discrimination training on stimulus generalization for human subjects. *Journal of Experimental Psychology*, *75*(4), 508.

Ghirlanda, S., & Enquist, M. (1998). Artificial neural networks as models of stimulus control. *Animal behaviour*, *56*(6), 1383–1389.

Hanson, H. M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, *58*(5), 321–334. doi: 10.1037/h0042606

Johnson, K., Flemming, E., & Wright, R. (1993). The hyperspace effect: Phonetic targets are hyperarticulated. *Language*, 505–528.

Lange, K., Kühn, S., & Filevich, E. (2015). " just another tool for online studies" (jatos): An easy solution for setup and management of web servers supporting online studies. *PloS one*, *10*(6), e0130834.

Lee, J., Lovibond, P., Hayes, B., & Lewandowsky, S. (2021). A mixture of experts in associative generalization. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 43).

Lee, J. C., Hayes, B. K., & Lovibond, P. F. (2018). Peak shift and rules in human generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *44*(12), 1955–1970. doi: 10.1037/xlm0000558

Lee, J. C., Mills, L., Hayes, B. K., & Livesey, E. J. (2021). Modelling generalisation gradients as augmented Gaussian functions. *Quarterly Journal of Experimental Psychology*, *74*(1), 106–121. doi: 10.1177/1747021820949470

Lee, M. D., & Wagenmakers, E.-J. (2014). *Bayesian cognitive modeling: A practical course*. Cambridge University Press. doi: 10.1017/CBO9781139087759

Lewis, M. B. (1999). Are caricatures special? evidence of peak shift in face recognition. *European Journal of Cognitive Psychology*, *11*(1), 105–117.

Lovibond, P. F., Lee, J. C., & Hayes, B. K. (2020). Stimulus discriminability and induction as independent components of generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *46*(6), 1106–1120. doi: 10.1037/xlm0000779

Lynn, S. K., Cnaani, J., & Papaj, D. R. (2005). Peak shift discrimination learning as a mechanism of signal evolution. *Evolution*, *59*(6), 1300–1305.

McLaren, I., & Mackintosh, N. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, *30*(3), 177–200. doi: 10.3758/BF03192828

Plummer, M. (2023). rjags: Bayesian graphical models using mcmc [Computer software manual]. Retrieved from `https://CRAN.R-project.org/package=rjags` (R package version 4-14)

Purtle, R. B. (1973). Peak shift: A review. *Psychological Bulletin*, *80*(5), 408–421. doi: 10.1037/h0035233

R Core Team. (2018). R: A language and environment for statistical computing [Computer software manual]. Vienna, Austria. Retrieved from `https://www.R-project.org/`

Schlegelmilch, R., Wills, A. J., & Lee, J. C. (2023, Sep). *Is generalization general? A comparison of predictive and category learning and the influence of response format and framing.* PsyArXiv. Retrieved from `psyarxiv.com/eakwz` doi: 10.31234/osf.io/eakwz

ten Cate, C., Verzijden, M. N., & Etman, E. (2006). Sexual imprinting can induce sexual preferences for exaggerated parental traits. *Current Biology*, *16*(11), 1128–1132.

Thomas, D. R. (1962). The effects of drive and discrimination training on stimulus generalization. *Journal of Experimental Psychology*, *64*(1), 24.

Thomas, D. R. (1993). A model for adaptation-level effects on stimulus generalization. *Psychological Review*, *100*(4), 658.

Thomas, D. R., Mood, K., Morrison, S., & Wiertelak, E. (1991). Peak shift revisited: a test of alternative interpretations. *Journal of Experimental Psychology: Animal Behavior Processes*, *17*(2), 130.

Thomas, D. R., Svinicki, M. D., & Vogt, J. (1973). Adaptation level as a factor in human discrimination learning and stimulus generalization. *Journal of Experimental Psychology*, *97*(2), 210.