



Perceptual variability: Implications for learning and generalization

Jonas Zaman^{1,2} · Anastasia Chalkia^{2,3} · Ann-Kathrin Zenses² · Antoine Selim Bilgin² · Tom Beckers^{2,3} · Bram Vervliet⁴ · Yannick Boddez^{2,5}

Published online: 31 August 2020
© The Psychonomic Society, Inc. 2020

Abstract

The generalization of learned behavior has been extensively investigated, but accounting for variance in generalized responding remains a challenge. Based on recent advances, we demonstrate that the inclusion of perceptual measures in generalization research may lead to a better understanding of both intra- and interindividual differences in generalization. We explore various ways through which perceptual variability can influence generalized responding. We investigate its impact on the ability to discriminate between stimuli and how similarity between stimuli may be variable, rather than fixed, because of it. Subsequently, we argue that perceptual variations can yield different learning experiences and that interindividual differences in generalized responding may be understood from this perspective. Finally, we point to the role of memory and decision-making within this context. Throughout this paper, we argue that accounting for perception in current generalization protocols will improve the precision of obtained generalization gradients and the ability to infer latent mechanisms. This can inspire future attempts to use generalization gradients as a (clinical) predictor or to relate them to individual traits and neural correlates and, ultimately, may lead to new theoretical and clinical insights.

Keywords Generalization · Perception · Discrimination · Learning · Memory · Decision-making · Conditioning

Introduction

Our environment constantly changes, both externally and internally, such that a (learning) situation is never replicated in its exact form. Hence, the capacity to generalize from previous

experiences is one of the most fundamental abilities of human and non-human animals, and the understanding of learning will never be complete without an understanding of generalization. Accordingly, attempts to characterize generalization have been performed across a wide range of species and behaviors (e.g., Ames & Yarczower, 1965; Armony, Servan-Schreiber, Romanski, Cohen, & LeDoux, 1997; Cohen, LeDoux, Armony, & Servan-Schreiber, 1997; De Hoz & Nelken, 2014; Ganz & Riesen, 1962; Solomon & Moore, 1975).

Associative learning paradigms have been the gold standard for the controlled investigation of generalization in a variety of situations. In classical conditioning, pairings of a cue (conditional stimulus, CS; e.g., a tone) with an outcome (unconditional stimulus, US; e.g., a shock) change responding to the former (e.g., it comes to elicit fear). In operant conditioning, the occurrence of a response (e.g., lever pressing) with an outcome (e.g., a food pellet) also results in a change in the former (e.g., an increase in lever pressing; De Houwer et al., 2013). Often, a stimulus is added that signals when the relation between the response and the outcome is present (e.g., a colored light). During a subsequent test phase (referred to as the generalization phase), multiple test stimuli (also called generalization stimuli; GSs) that differ from the initial

Anastasia Chalkia, Ann-Kathrin Zenses and Antoine Selim Bilgin shared second co-authorship. As all authors contributed equally author order was decided by random permutation.

✉ Jonas Zaman
Jonas.zaman@kuleuven.be; <https://ppw.kuleuven.be/ogp>

¹ Health Psychology, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, Box 3726, 3000 Leuven, Belgium

² Center for the Psychology of Learning and Experimental Psychopathology, Faculty of Psychology and Educational Sciences, KU Leuven, Tiensestraat 102, Box 3712, 3000 Leuven, Belgium

³ Leuven Brain Institute, KU Leuven, Leuven, Belgium

⁴ Laboratory for Biological Psychology, KU Leuven, Tiensestraat 102, Box 3714 Leuven, Belgium

⁵ Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium

stimulus, are presented. Most often, the test stimuli are created by changing stimulus features within a specific physical dimension, such as size, frequency, or hue. However, variations across other dimensions, like category membership, have also been used successfully (Dunsmoor, White, & LaBar, 2011). The outcome of interest is the conditioned response (CR) elicited by the range of test stimuli, and assessed by a variety of measures (e.g., lever presses, US expectancy ratings, freezing, skin conductance responses, startle eyeblink potentiation, pupil dilation).

Thus, during a typical generalization phase, responses to the repeated presentation of a set of stimuli are recorded, with the stimulus set most often comprising minor differences of a specific stimulus feature. Responses are averaged per stimulus and plotted along the physical dimension on which the stimuli differ, creating a gradient for each subject (Fig. 1). These individual gradients are then averaged into a group gradient yielding the well-established bell-shaped gradient. Its frequent observation in the laboratory has fostered the idea that this pattern is universal among individuals (Shepard, 1987). Although on a group level striking similarities are found across a wide range of experiments (Ghirlanda & Enquist, 2003; Mednick & Freedman, 1960), this is hardly the case for individual gradients (Fig. 1). The extent of response variance underlying average group gradients is often underestimated, as the majority of generalization studies only focus on the latter. In addition, the theoretical assumption underlying the expression of behavior in relation to physical stimulus features is one of perception as a static representation of the physical properties of the world (Struyf, Zaman, Vervliet, & Van Diest, 2015). At best, perceptual variability is considered a mere source of noise (Atkinson & Estes, 1963; McLaren & Mackintosh, 2002) that will be canceled out via averaging across trials and subjects. It has been demonstrated that cognitive models strongly differ in their ability to account for group-averaged data compared to individual data. The model that best accounts for group-averaged data is not necessarily the model that can best explain data on an individual

level (Maddox, 1999). Via averaging, patterns that do not exist on an individual level may appear (Estes, 1956; Hayes, 1953; Merrell, 1931; Sidman, 1952; Struyf et al., 2015). Thus, we maintain that, as a consequence of this practice, mechanisms underlying generalized responding remain poorly understood (see Box 1). Given that the concept of generalization has been implicated in many forms of psychopathology, the further refinement of existing experimental protocols is not only of theoretical but also of great clinical importance. Such improvements may enable researchers and clinicians to scrutinize and identify the various mechanisms that may contribute to generalized responding, as we will demonstrate in the remainder of the paper.

Similarity

Originating in the early 1900s, the procedures to study generalized responses led to the discovery of the by now well-established generalization gradient, where response strength (or response probability) typically varies as a function of physical difference (Spence, 1937). The larger the physical difference between a test stimulus and the initial stimulus, the less likely the former is to elicit a CR (or the less intense the CR is) (Fig. 1). This observed inverse relationship has inspired the notion that generalized responding is a function of the similarity between the initial learning situation and the novel situation (Ghirlanda, 2015). Given that the brain cannot directly access the physical properties of a stimulus, it must rely on input from the senses. Consequently, all cognitive generalization theories include mental representations of the stimuli from our environment.

Depending on how they represent stimuli, the quantitative modeling of similarity will be different, but its assumed role in generalization is similar across theories (Ghirlanda, 2002, 2015). For instance, stimulus sampling (Atkinson & Estes, 1963) and elemental models (Blough, 1975; McLaren & Mackintosh, 2000, 2002) postulate that

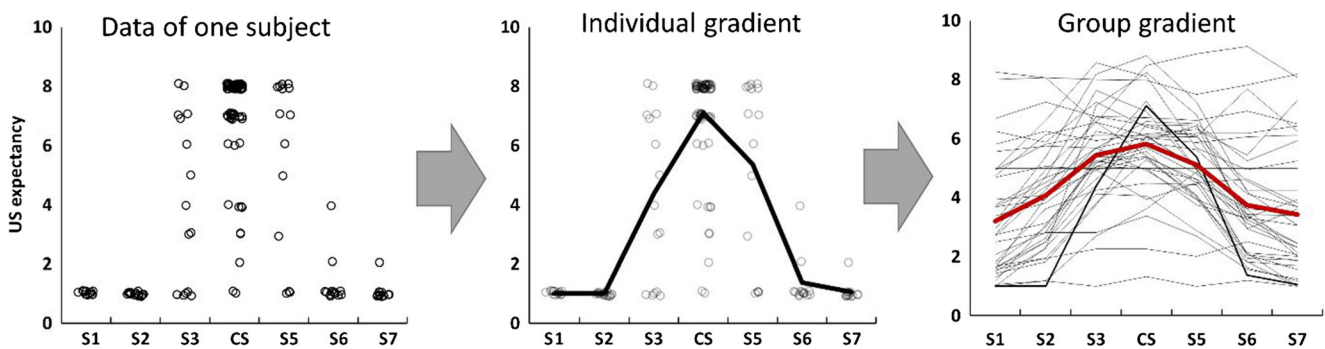


Fig. 1 Illustration of the variation in conditioned responses underlying individual gradients and the average group gradient. **Left:** The raw data of a subject from a generalization phase. **Middle:** The individual gradient through averaging responses per stimulus. **Right:** The group gradient

(red) by averaging individual gradients across the stimulus dimension. Data reused with permission from Zaman, Struyf, Ceulemans, Beckers, & Vervliet, 2019b. CS conditioned stimulus, S test stimulus, US unconditioned stimulus

the mental representation of a stimulus consists of a set of activated units that correspond to a sample of all the physical elements of that stimulus. Therefore, a physically similar stimulus will be represented by a partially overlapping set of elements. Likewise, neural models use activation patterns of sensory neurons to represent stimuli as distributions across these neural units (Enquist & Ghirlanda, 2005; Ghirlanda & Enquist, 2007). If there is more overlap between two stimulus distributions, their similarity will be higher, and the trained response will generalize more from one to the other. All of these models use a Gaussian distribution across units to represent a stimulus (or its activation pattern). However, few are currently able to provide a theoretical rationale for what determines the location of the peak, the width of these distributions, or how they may change due to previous experience. Rather, these are unknown parameters inferred post hoc from behavioral data (Ghirlanda, 2002). Another custom way of representing stimuli, especially in the context of categorization models, is to think of them as points in a multidimensional psychological space, with each dimension representing a physical feature (e.g., color, size) (Jäkel, Schölkopf, & Wichmann, 2008; Kruschke, 1992; Nosofsky, 1986). Here, the similarity between two stimuli is determined by their distance in psychological space (Shepard, 1987; Shepard, 1958), which, again, is assumed to be held constant.

Although the notion of perceptual variability has been (explicitly or implicitly) acknowledged by some authors, only few have considered in depth the implications of perceptual variability on stimulus representations, and subsequently, on learning or generalization (Ashby & Maddox, 1993; Ashby & Townsend, 1986; Atkinson & Estes, 1963; Enquist & Ghirlanda, 2005; Estes & Burke, 1953; McLaren & Mackintosh, 2002). Most generalization theories assume that the representation of a stimulus is veridical (i.e., accurately corresponding to the physical reality). As such, they reduce similarity to a physical dimension and, in essence, model behavior solely in relation to the physical features of the stimuli as presented during testing. Thus, they assume perception as a static process unaffected by previous experience, despite research demonstrating the plasticity of the perceptual system (Gibson, 1969; Gilbert, Sigman, & Crist, 2001; Sagi, 2011). However, false assumptions about stimulus representations may lead to incorrect inferences regarding underlying cognitive processes (Maddox, 1999).

When we inspect the amount of variation in responses underlying both an individual gradient and a group gradient (Fig. 1), it becomes apparent that modeling responding solely based on the physical features of the presented stimuli fails to account for a large amount of response variance. However, as most studies only report

average group gradients, the extent of trial-by-trial (intraindividual) and interindividual variability in responding to the same physical stimulus remains understated. As such, it may seem that there is little necessity to extend or revise contemporary generalization protocols and models. However, in the remainder of the paper, we demonstrate how inclusion of perceptual measures in generalization research may lead to a better understanding of both types of response variance and, ultimately, to new theoretical and clinical insights. A more explicit focus on perception could be of both heuristic and predictive value to researchers as: (1) response *noise* due to perceptual variability would be accounted for; (2) model parameters that are currently inferred post hoc from gradients in generalized responding could be estimated based on perceptual data; and (3) the incorporation of perceptual models into generalization research would allow predictions of why, when, and how perception varies, thereby affecting generalization. Furthermore, even if researchers are not interested in perception and consider it a purely confounding factor that they aim to average out, we argue that accounting for it in current generalization protocols will improve the precision of obtained generalization gradients. This can inspire future attempts to use generalization gradients as a (clinical) predictor or to relate them to individual traits and neural correlates.

Box 1

The term generalization is used in many different disciplines and has been used both as an explanandum (i.e., that which needs to be explained) and the explanans (i.e., that which is used to explain). As an explanandum, the term generalization is used to refer to how various situations can elicit similar behavior without making any assumptions regarding what drives the behavior. Thus, whether a rat freezes to a test tone simply because it cannot perceptually discriminate the test tone from the trained tone is an open question when generalization is used as an explanandum. However, as the explanans, it is used to explain why behavior is similar across situations. For example, Jäkel et al. (2008) wrote the following: “As a theoretical construct, generalization refers to a covert process that leads a subject to respond to a new stimulus in the same way as to a previously learned stimulus, despite the ability of the subject to tell the stimuli apart” (Jäkel et al., 2008, p. 258). In our rat example, it then becomes imperative to ensure that the rat is able to perceptually discriminate between the tones, before inferring that generalization drove its behavior. Unfortunately, in current experimental protocols, it remains difficult to assess whether response gradients reflect a generalization process, some degree of confusion in a memory process, or perceptual indiscriminability (Jäkel et al., 2008; Struyf et al., 2015). Furthermore, generalization has been used interchangeably as an explanandum and the explanans, thereby fostering theoretical ambiguity. The interchangeable use of a concept is not without danger. It has the potential to result in a tautology: generalization was observed because of generalization. As many clinical applications are motivated by empirical generalization findings, it is important to be aware of the existing limitations of current experimental protocols. Differences in response gradients do not necessarily reflect differences in generalization as a process, as we will demonstrate below.

Perceptual variability

An abundance of research has demonstrated that perception of the same stimulus is not static. In almost all research that investigates how humans perceive stimuli, using categorizations, identifications, judgments, or adjustment tasks, responses vary within and between individuals (e.g., Hoskin et al., 2019; Huang & Sekuler, 2010; Jones, Love, & Maddox, 2006; Li, Howard, Parrish, & Gottfried, 2008; Ons, De Baene, & Wagemans, 2011; Petzschner & Glasauer, 2011; Samaey, Wagemans, & Moors, 2020; Schechtman, Laufer, & Paz, 2010; Tenenbaum & Griffiths, 2001; Zaman, Wiech, & Vlaeyen, 2019c). Ashby and Lee (1993) even suggested that the success of perceptual theories depends on their ability to account for what they called “the axiom of perceptual variability,” stating that variations in perceptual representations are fundamental to perception research. Furthermore, this perceptual variability is not mere noise originating from our senses, as assumed by most theories concerning generalization, but can also be explained by various perceptual models (e.g., Davis & Love, 2010; Feldman, Griffiths, & Morgan, 2009; Friston & Kiebel, 2009; Hoskin et al., 2019; Petzschner, Glasauer, & Stephan, 2015; Press, Kok, & Yon, 2020; Tenenbaum & Griffiths, 2001). Many different effects on perception have been observed in the laboratory, but those most relevant to generalization research are, among others, effects of preceding stimuli (e.g., Chambers & Pressnitzer, 2014; Jones et al., 2006; Petzschner & Glasauer, 2011), expectations (e.g., Wiech et al., 2014; Zaman et al., 2019c), stimulus repetition (e.g., Dong, Gao, Lv, & Bao, 2016; Muenssinger et al., 2013; Pérez-González & Malmierca, 2014), the range of presented stimuli (e.g., Kowal, 1993; Petzschner & Glasauer, 2011), and associative learning at large (e.g., Asutay & Västfjäll, 2012; Resnik, Sobel, & Paz, 2011; Zaman et al., 2018). For instance, Petzschner et al. (2015) demonstrated that size estimates were over- or underestimated depending on the range of presented stimuli and were able to model such effects using a Bayesian perception model (see below). In addition, the authors were able to account for effects of presentation order via the incorporation of preceding stimulus presentations in their model. Similarly, others have demonstrated serial dependency in perception, in that stimuli presented during previous trials affected the perception of the current stimulus (e.g., Cicchini, Mikellidou, & Burr, 2017; Jones et al., 2006; Jones & Sieck, 2003).

Nonetheless, perception and generalization research have remained relatively separate fields. Generalization researchers often make use of custom-made stimulus sets for which no database with similarity ratings exists. Furthermore, the majority of studies only measure conditioned responses without a

measure of the trial-by-trial stimulus perception making it difficult to assess the extent of perceptual variability. Only a handful of studies report perceptual data for stimulus sets used to study generalization (Holt et al., 2014; Lovibond, Lee, & Hayes, 2019; Onat & Büchel, 2015; Struyf, Zaman, Hermans, & Vervliet, 2017; Zaman, Ceulemans, Hermans, & Beckers, 2019a; Zaman, Struyf, et al., 2019b). A stimulus set that has been frequently used to study the generalization of a trained fear response, both in healthy volunteers and clinical populations, consists of ten circles, ranging from 5.08 cm to 11.49 cm, with steps of 0.72 cm in between (e.g., Lange et al., 2017; Lissek et al., 2008, 2014, 2010; Timoco-González et al., 2015; Vervliet, Iberico, Vervoort, & Baeyens, 2011).

In a recent series of studies, we investigated the extent of perceptual variability in this stimulus set (Struyf et al., 2017; Zaman, Ceulemans, et al., 2019a; Zaman, Struyf, et al., 2019b; Zaman, Struyf, Ceulemans, Vervliet, & Beckers, 2020). Perception was assessed using a categorization task. Following a fear-conditioning phase in which one circle was paired with an aversive outcome, subjects were instructed to categorize each newly presented circle as either the same stimulus that was presented during the previous (conditioning) phase or as a different stimulus. We found not only numerous perceptual errors for the test stimuli (mean error = 41.35%, SD = 17% – Zaman, Struyf, et al., 2019b; mean error = 32%, SD = 32.7% – Zaman, Ceulemans, et al., 2019a) but also multiple errors on trials on which the training circle (CS) was presented (mean error = 34.52%, SD = 29% – Zaman, Struyf, et al., 2019b; mean error = 44%, SD 35.2% – Zaman, Ceulemans, et al., 2019a). These findings suggest a great extent of perceptual variability both between and within individuals for this stimulus set (Fig. 2). The repeated presentation of the same stimulus seemed to have led to different percepts within a subject (intraindividual variability). In addition, large interindividual differences were observed in the overall extent of committed errors, as well as in the distribution of errors across the stimulus spectrum. In Fig. 2, for almost half of the subjects, test stimuli at one end of the test dimension (either the smallest or largest circles) had higher probabilities of being (mis)perceived as the trained stimulus compared to the actual stimulus used during training. The other half was relatively accurate. They most often correctly categorized the training stimulus. Committed perceptual errors were more likely for test stimuli near the training stimulus than for test stimuli that were more different, thus yielding a perceptual gradient comparable to the bell-shaped generalization gradient (Fig. 2). This strong similarity suggests that perceptual errors may indeed shape a generalization gradient to a large extent, or alternatively, that both are driven by the same latent process (although we believe the latter is less likely for reasons outlined below, see Box 2).

In another study, perception was measured using an estimation task with the same stimulus set. On every trial, subjects

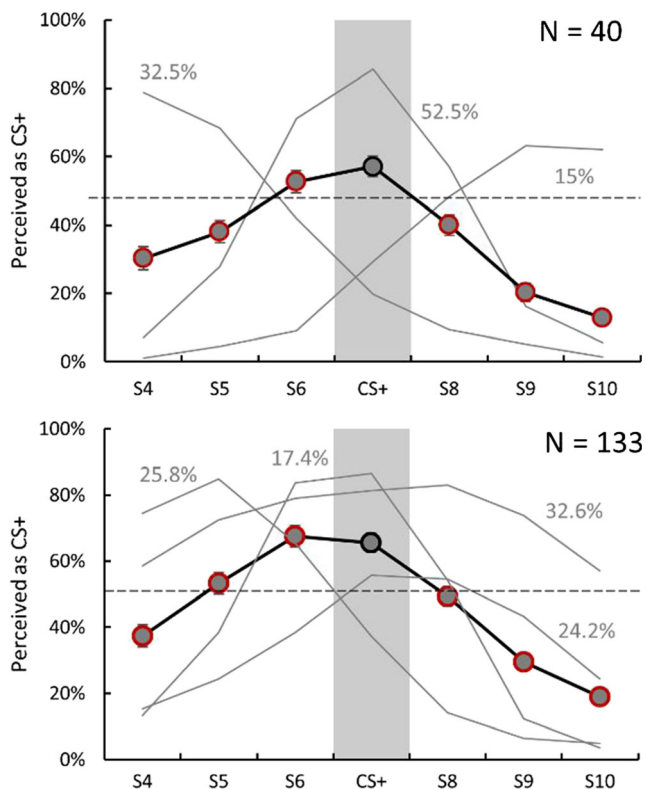


Fig. 2 Data on perceptual accuracy during generalization testing after a simple conditioning procedure (**top**: Zaman, Struyf, et al., 2019b; **bottom**: Zaman, Ceulemans, et al., 2019a). The black line is the averaged probability gradient that stimuli are (mis)perceived as the stimulus from the training phase. The red circles indicate perceptual errors, the dotted line indicates chance performance. The gray lines represent identified subgroups of inter-individual differences in perceptual accuracy using cluster analyses, and their corresponding percentages show the number of subjects within each subgroup. CS+ conditioned stimulus, S test stimulus

were prompted to estimate the size of the presented circle using a visual analog scale (Zenses, Lee, & Zaman, 2020 preprint). We found large trial-by-trial variability in size estimates for each stimulus within subjects (average subject SD = 6.33 cm), with the average subject SD being almost tenfold of the actual physical difference between stimuli (which was 0.72 cm). Clearly, the impact of such perceptual variability on generalization depends on the physical differences between the presented stimuli. However, as already mentioned, most generalization research uses stimuli sets with small physical differences. The ratio between the extent of perceptual variability and the size of the manipulated physical differences determines the impact of the former. If the physical difference between stimuli is large, perceptual variability is unlikely to affect perceptual discrimination. On the other hand, if physical differences are small, perceptual variability may lead to indiscriminability, resulting in many perceptual mistakes. Alternatively, as we discuss later, it could affect similarity more gradually.

Perceptual discrimination and generalization

Although the issue of perceptual discrimination has been intensely debated since the early days of generalization research, it remains unresolved. Lashley and Wade (1946) proposed the inverse hypothesis stating that generalization occurs merely due to a failure in perceptual discrimination, where the inability to detect that a test stimulus is distinct from the CS drives a generalization gradient. According to these authors, the phenomenon of generalization would then be a mere result of (inaccurate) perception. Despite the wide-ranging clinical and theoretical implications of this alternative account, relatively few studies have attempted to investigate the relationship between perceptual discrimination and generalized responding (De Hoz & Nelken, 2014; Guttman & Kalish, 1956; Holt et al., 2014; Honig & Urcuioli, 1981; Kalish, 1958; Lovibond et al., 2019). For example, Guttman and Kalish (1956) obtained gradients in conditioned responding across a light wave spectrum where identical physical changes in hue were assumed to result in different levels of perceived change in color. Despite expected differences in perceptual sensitivity [based on the work of Hamilton and Coleman (1933)], depending on the location of the CS across the spectrum, similar gradients were obtained. Follow-up work, however, demonstrated that Guttman and Kalish (1956) selected stimuli based on incorrect psychometric functions and that the obtained gradients did relate to variations in perceptual sensitivity (Blough, 1972; Honig & Urcuioli, 1981). In another study, Holt et al. (2014) created GSs based on their subject's psychometric functions, obtained before the experiment using a two-forced choice task (same vs. different). They found no evidence of generalization for stimuli located above the just-noticeable difference (JND) threshold. In contrast, Onat and Buchel (2015) found fear-generalization gradients across a range of discriminable faces. The stimuli's perceptual organization was inferred from a two-alternative forced-choice (2-AFC) similarity judgment between two face pairs. Importantly, however, performance during a 2-AFC task, where stimuli are presented simultaneously or separated by a short interval, might substantially differ from the ability to differentiate a separately presented test stimulus from a CS retrieved from memory (Struyf et al., 2015). Furthermore, selecting stimuli based on prior discrimination thresholds and responses per stimulus is still an approach that regards perception as a static process unaffected by previous experience. Yet, in the next section we demonstrate that perception is malleable by learning, with established effects of conditioning on discrimination acuity, among others.

Conditioning effects on perception

Of particular interest is that the same conditioning protocols used to study generalization also affect perception (Asutay & Västfjäll, 2012; Stolarova, Keil, & Moratti, 2006), and more precisely, the ability to discriminate a test stimulus from the CS (Aizenberg & Geffen, 2013; Laufer & Paz, 2012; Resnik et al., 2011; Schechtman et al., 2010). For example, in a study by Schechtman et al. (2010), three tones were associated with either monetary gain, monetary loss, or no outcome during a training phase. In a subsequent test phase, similar but not identical tones to the previously used tones were also presented. Subjects had to indicate whether the tone was novel or was one of the three tones presented during training. Importantly, during this phase, reinforcement (monetary gain) was dependent on the correctness of the perceptual classification and no longer on the presented stimulus, although no feedback was provided upon responding. More perceptual errors were made for tones surrounding the CS tones (previously associated with monetary gain or loss) compared to the control tone (paired to no outcome), suggesting that learning modulated subjects' ability to perceive differences. For instance, test tones that differed 60 and 100 Hz from the initially control tone had a probability of being mistaken for that tone that was close to zero. For equidistant test stimuli to the CS tone paired with monetary loss, the probability of being mistaken for the CS increased approximately to 70% and 55%. Resnik et al. (2011) extended those findings by showing that discrimination thresholds increased by 20% after an aversive conditioning procedure, indicative of a deteriorated performance in discriminability.

Others have demonstrated improvements in perceptual discriminability after an atypical differential aversive conditioning protocol where initially indiscriminable odor enantiomers (mirror image molecules) were used as CSs (Åhs, Miller, Gordon, & Lundström, 2013; Li et al., 2008). In a series of experiments, Aizenberg and Geffen (2013) demonstrated that effects of fear conditioning on tone discrimination can occur in either direction and that this depends on the presence of another cue during learning (i.e., differential learning) and the physical distance between both cues. Differential conditioning improved perceptual discrimination only with physically proximal cues, while physically distant cues decreased discrimination acuity of tones around the reinforced tone. Simple conditioning yielded the worst perceptual discrimination (Aizenberg & Geffen, 2013).

Importantly, these effects were not limited to changes in performance on perceptual tasks, but also involved neural alterations at the level of the primary sensory cortices. These included tuning profiles of single neurons and activation patterns, indicative of alterations in the processing of sensory input (Aizenberg & Geffen, 2013; Chen, Barnes, & Wilson, 2011; Diesch & Flor, 2007; Kass, Rosenthal, Pottackal, &

McGann, 2013; Li et al., 2008). For instance, the perceptual improvements in the Li et al. (2008) study were paralleled by differential activation patterns in the primary olfactory cortex. Compared to before conditioning, spatial activity patterns diverged for the enantiomer pair that was differentially paired to the aversive outcome, whereas patterns for the control enantiomer pair remained stable and highly correlated. Aizenberg and Geffen (2013) demonstrated, using reversible pharmacological inactivation, that these perceptual changes occurred at the level of the auditory cortex. Finally, conditioning-induced changes in discrimination acuity and sensory processing demonstrate that discrimination performance assessed prior to conditioning differs from discrimination performance during generalization testing.

All of the abovementioned studies show that the learning procedures used to study generalization also affect perception. Accordingly, this should be taken into account when interpreting the results of generalization experiments. For example, patients with generalized anxiety disorder (GAD) exhibited a reduced ability to perceive test stimuli as different from the trained stimulus only after conditioning (using both positive and aversive USs) compared to controls (Laufer, Israeli, & Paz, 2016). This finding allows for an alternative explanation of reported differences in generalization gradients between anxiety patients and controls (Lissek & Grillon, 2010; but see Tinoco-González et al., 2015). Instead of attributing them to a risk aversion bias (Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015), elevated generalization gradients could be explained by more perceptual errors (the perception of a test stimulus as a CS) in patients compared to controls. Wider generalization gradients in patients could thus merely reflect larger proportions of perceptual errors on test trials, with test stimuli perceived as the initially trained stimulus eliciting stronger conditioned responses compared to when the same stimulus is correctly perceived.

Perceptual variability and generalized responding

Perception can be measured using different types of tasks, and its theorized impact on generalized responding will depend on it. We will start with research that approaches perception from a categorical perspective. In the context of generalization, in its simplest form, the perceptual task subjects face is to decide whether the presented stimulus is the initially trained stimulus or a different one. In other words, one is either able to discriminate the test stimulus on a given trial accurately or not.

In a series of studies, we recently demonstrated that during a generalization protocol, stimuli are far from being consistently or accurately discriminated (Struyf et al., 2017; Zaman, Ceulemans, et al., 2019a; Zaman, Struyf, et al., 2019b; Zaman et al., 2020). Findings of these studies

indicated that the incorrect perception of a test stimulus (i.e., subjects mistakenly perceived it as the trained stimulus on that trial) was associated with stronger self-reported and psychophysiological conditioned responses, while the correct identification of the same stimulus elicited a much lower CR (Fig. 3). Many mistakes were also made regarding the trained stimulus (Struyf et al., 2017; Zaman, Ceulemans, et al., 2019a; Zaman, Struyf, et al., 2019b; Zaman et al., 2020). When subjects incorrectly perceived this stimulus as a GS, conditioned responding was attenuated compared to when their perception of the stimulus was correct. Apart from this intraindividual variability, we also found large inter-individual differences in perception that affected the shape of the response gradient in different manners (Zaman, Ceulemans, et al., 2019a; Zaman, Struyf, et al., 2019b; Zaman et al., 2020). More precisely, the distribution of perceptual errors differed greatly between individuals. These differences in perceptual accuracy predicted differences in the shape of the generalization gradient between individuals. Apart from this direct effect, perceptual errors on CS trials (which were reinforced by the US) led to different experienced stimulus-outcome probabilities that affected the gradient. The more a subject experienced, due to committed perceptual errors, that both the trained and test stimuli had similar reinforcement probabilities, the further the differential effect of perception dissolved and the flatter the gradient became (Fig. 3).

In a recent follow-up study (Zaman et al., 2020), a differential instead of a simple conditioning protocol was used. During the learning phase, another different-sized circle (i.e., a CS-), predictive of the absence of the aversive outcome, was also included. Its inclusion during training yielded two interesting, and previously documented, phenomena, called the area and peak shift, which have played a central role in the development of generalization theories. The peak shift relates to the observation that after differential learning, the peak in the response gradient shifts away from the reinforced stimulus (CS+) in the direction opposite to the location of the non-reinforced stimulus (CS-) (Hanson, 1957; Hanson, 1959; Purtle, 1973). The area shift refers to the finding that the symmetrical gradient typically found after simple conditioning (Honig & Urcuioli, 1981) changes into a skewed gradient in the direction away from the CS-, with differential conditioning. Interestingly, we also observed an area and peak shift for the perceptual gradient based on the categorization task (Fig. 4). That is, not the actual CS+ but an adjacent test stimulus had the highest probability of being perceived as the CS+ (80.64%, SE = 2.02%).

A shift in perception has been found previously in perceptual categorization research (Davis & Love, 2010). These authors proposed that category averages (i.e., prototypes) are not determined solely by the statistical properties of the perceived stimuli but can shift depending on situational demands in the direction opposite to the contrasted category (Davis & Love,

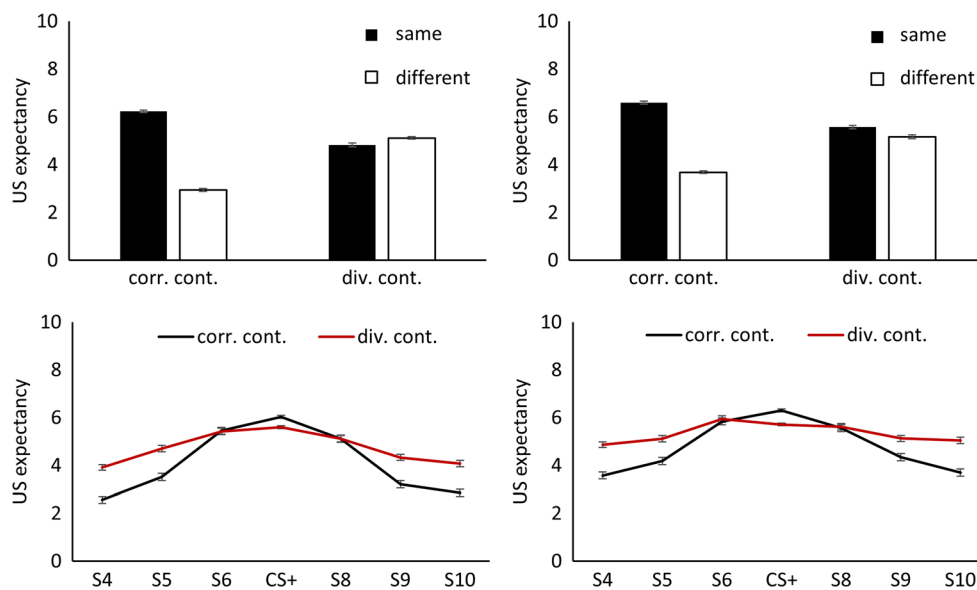


Fig. 3 Data on self-reported US expectancy during generalization (left: Zaman, Struyf, et al., 2019b, right: Zaman, Ceulemans, et al., 2019a). **Top:** Average US expectancy for when a stimulus was perceived as the same stimulus of the previous (training) phase (i.e., the CS) or as different one (i.e., GS). Depending on the patterns of perceptual errors, experienced CS-US and GS-US contingencies could differ between subjects. Subjects either experienced contingencies that corresponded to the

objective contingencies (CS-US > GS-US) or experienced divergent contingencies in that both the trained and test stimuli had similar reinforcement probabilities (CS-US = GS-US). **Bottom:** Averaged US expectancy gradients for both groups. *corr. cont.* corresponding contingency, *div. cont.* divergent contingency, *CS+* conditioned stimulus, *US* unconditioned stimulus

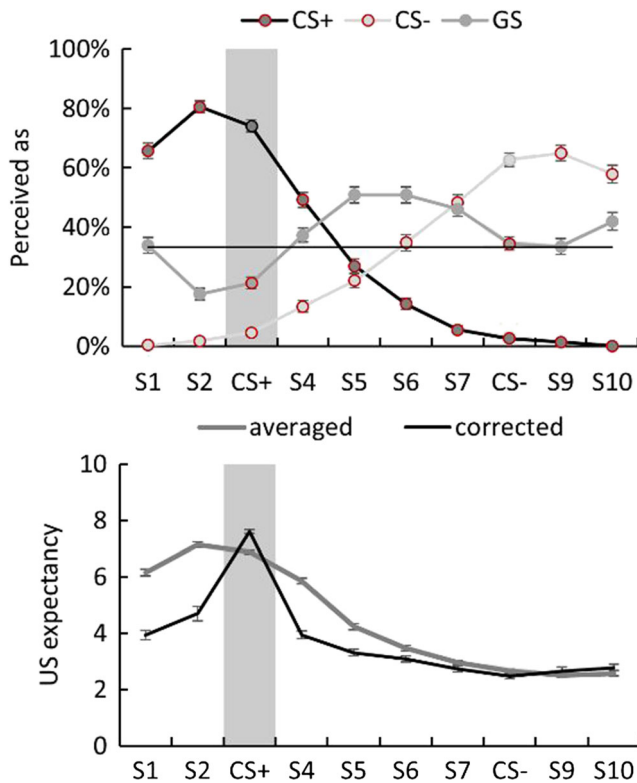


Fig. 4 Data on perceptual accuracy and self-reported US expectancy during generalization after differential conditioning (Zaman et al., 2020). **Top:** Average probabilities across the stimulus spectrum for the different perceptual response categories. The red circles indicate perceptual errors and the black line indicates chance performance. **Bottom:** Average US expectancy ratings based on all trials (averaged gradient) and on trials with accurate categorization performance only (corrected gradient). CS+ reinforced conditioned stimulus, CS- unreinforced conditioned stimulus, S generalization stimulus, US unconditioned stimulus

2010). It remains to be determined if and to which extent a perceptual peak shift is driven by similar mechanisms. Furthermore, we found that correcting for such perceptual errors resulted in the disappearance of the peak shift and a strong reduction of the area shift in the US expectancy gradient (Fig. 4). In a series of experiments, Lovibond and colleagues (Lovibond et al., 2019) only found a peak shift in their generalization gradient after the inclusion of additional test stimuli that had high probabilities of being mistaken for the CS+, due to their close resemblance to it. These findings suggest that stimulus indiscriminability may play a key role in such specific generalization phenomena. Furthermore, a comparison of the perceptual accuracies between this and our previous studies revealed that fewer perceptual errors were made during testing on CS+ trials after differential compared to simple conditioning [accuracy CS+ differential: 74.14% (SE = 2.02%) vs. simple: 65.48% (SE = 2.5%)] (Zaman, Ceulemans, et al., 2019a; Zaman et al., 2020). This is in line with previously reported effects of learning on perception (Aizenberg & Geffen, 2013).

This research demonstrated that perception strongly affects the degree of conditioned responding in certain experimental procedures used to investigate generalization. Thus, a generalization gradient in experiments that rely on the use of test stimuli that differ on a physical continuum (e.g., size), might in part emerge due to an underlying probability gradient of perceptual errors (Struyf et al., 2015), with perceptual errors altering the shape of the obtained gradient (see Fig. 4 and 5). In this section, we only explored and demonstrated the impact of stimulus indiscriminability. Next, we illustrate how perceptual variability may influence similarity.

Depending on how generalization models formalize stimulus representations, perceptual variability could be implemented in different manners. If perception rather than physical features determines the coordinates of the stimulus representations in a multidimensional psychological space, perceptual variation will impact distance-based similarity estimates (Shepard, 1987; Shepard, 1958; Shepard & Chang, 1963). Rather than similarity being fixed between two stimulus representations, it would become variable as a function of perception. Depending on the direction of the perceptual change, the similarity between the same stimulus pair will decrease or increase with analog effects on the extent of generalized

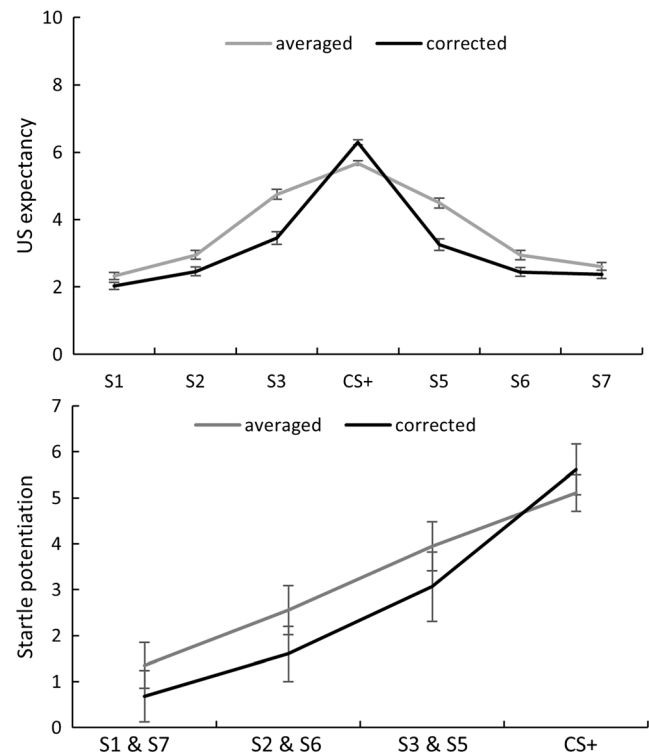


Fig. 5 Data on self-reported US expectancy and startle eye blink potentiation during generalization (Zaman, Struyf, et al., 2019b). Average US expectancy ratings (**top**) and startle data (**bottom**) based on all trials (gray line, averaged gradient) and on trials with accurate categorization performance only (black line, corrected gradient). CS+ reinforced conditioned stimulus, S generalization stimulus, US unconditioned stimulus

responding. In generalization models that consider stimuli as probability distributions of activated neural (or elemental) units, perceptual variability would translate to shifts of the peak, or it may affect the width of the distribution of activated units. Again, depending on the shift, the overlap between the distributions would affect similarity between the same stimulus pair allowing for variation in generalized responding to the same test stimulus (currently ignoring other effects like continuous learning, e.g., extinction).

The concept of probability distributions is also central to current notions of perception as Bayesian inference (e.g., Clark, 2013; Hoskin et al., 2019; Petzschner et al., 2015; Press et al., 2020). According to this view, perception is the computation of the most likely physical source given certain sensory input and a priori beliefs. It assumes that a physical source (i.e., a stimulus) transforms into a distribution of (potential) sensory input. Certain sensory inputs will be more likely than others for a given physical source. This distribution (called the likelihood) expresses the probability of sensory input given a physical cause (Fig. 6). At this point, the idea is very similar to many generalization theories. However, this sensory input is not interpreted in isolation. Observers often have expectations about which sensations are more likely. Prior information based on regularities in the environment results in a distribution of expected sensations called the prior. It is the combination of sensory input (likelihood) with prior information (prior) that results in a posterior distribution of potential sensations. The impact of the likelihood and the prior

on the posterior depends on the width of their distribution (precision). A posterior will be more driven by sensory input characterized by high compared to low precision. As the posterior on a given trial will act as a prior on the subsequent trial, the brain actively constructs and continuously updates a generative model based on sensory input (see Fig. 6).

The main difference compared to generalization research is that these theories allow for the updating of the representational stimulus distributions based on previous experience. The extension, for instance, of generalization theories with the notion of perception as Bayesian inference could not only enable generalization theories to incorporate perceptual variability but also model effects of previous experience, learning, and expectations on perception. This would allow researchers to model regularities in how the repeated presentation of the same stimulus affects representational distributions, and, consequently, how the similarity between the same pair of stimuli would differ across trials. For example, in the beginning of a generalization test, the prior is expected to be centered around the CS location, because, based on the subject’s experience, this stimulus is most likely (as it was the only presented stimulus during learning). Across generalization testing, different stimuli are presented. Hence, the prior should widen, reflecting the subject’s experience that many different stimuli now can be expected. As a consequence, stimuli should be more likely to be perceived as the CS during initial test trials compared to later trials as the *biasing* effect of the prior disappears due to its widening, which we indeed found in our

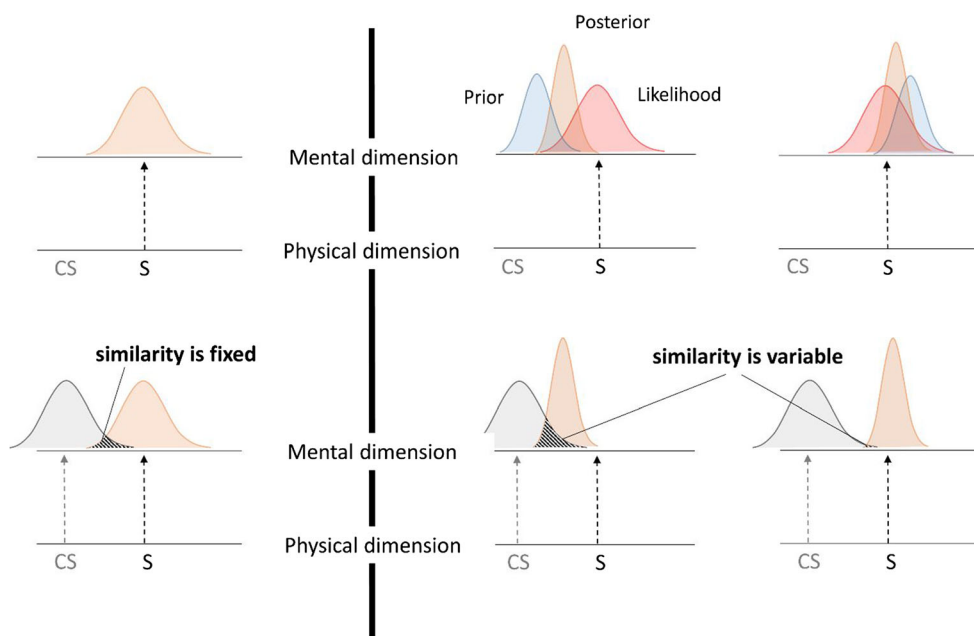


Fig. 6 Schematic representation of stimulus representation and similarity. **Left:** Veridical stimulus representation (implicit in generalization analyses). A stimulus is represented by an invariant distribution. It follows that similarity defined by the overlap in distribution between a given stimulus pair is fixed. **Right:** Bayesian perception model. Perception is an inferential process whereby the brain makes inferences

about the physical world (S) based on current sensory input (likelihood) and a priori beliefs (based on past expectations, previous trials, etc.). The same stimulus presented repeatedly will not necessarily result in the posterior distribution. Hence, the similarity between a stimulus pair can vary depending on the shape and location of the posterior. CS conditioned stimulus, S test stimulus

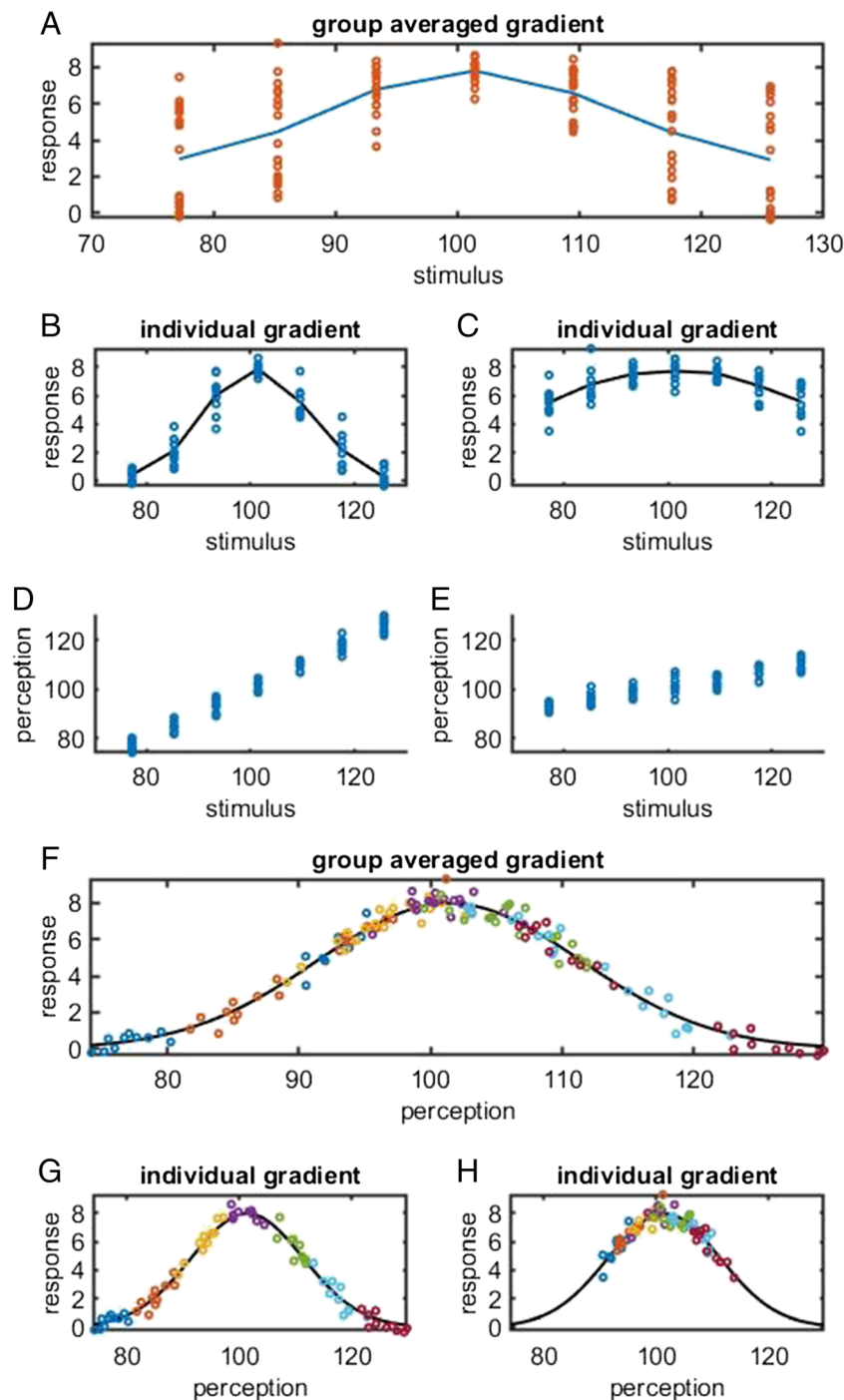


Fig. 7 Simulated data set. (A) The averaged group gradient and underlying responses (dots) across a physical dimension. (B, C) Individual response gradients. (D, E) Scatterplot between stimulus size

and perceived size. (F) The averaged group gradient across a perceptual dimension (dot color denotes the different stimuli). (G, H) Individual response gradients across a perceptual dimension

studies (unpublished analyses, data available at <https://osf.io/t8u92/>).

In Fig. 7, we simulated perceptual variability for two fictitious subjects by drawing their stimulus perception from normal distributions with means either equal to the actual stimulus size or closer to each other, reflecting difficulties to differentiate between stimuli. The standard deviation did not differ across

stimulus distributions nor between subjects (and was set to the previously found SD of 6.33 cm). We used a Gaussian kernel function based on stimulus perception plus some random noise to generate generalized responses. The graphs (Fig. 7) are based on ten trials per stimulus per subject. The top graph is the averaged group gradient as typically done, plotting average responses per stimulus. A comparison between the averaged

individual gradients (Fig. 7 B and C), would suggest *overgeneralization* in the second subject as the gradient is much flatter. However, perceptual estimates for the different stimuli were much closer to each other for this subject compared to the other (Fig. 7 D and E). Finally, we plotted responses along the perceptual dimension (where color denotes the physical stimulus). Now generalization gradients are alike between individuals, suggestive of similar generalization tendencies (Fig. 7 G, H). As such, gradients of conditioned responding across a range of similar physical stimuli, without insight into the perception of the stimulus eliciting the behavior, will provide a limited understanding of the underlying mechanism and remain at best purely descriptive. Note that apart from causing interindividual variability, intraindividual response variability is also reduced when perceptual variations are incorporated into generalization gradients.

Perceptual variability during learning

Until now, we have solely focused on perceptual variability during the assessment of generalized responding. However, a crucial premise to observe generalization of a learned response is that some sort of learning precedes it. Typically, subjects learn about co-occurrences between stimuli or behaviors and outcomes. A bulk of research has demonstrated that manipulating various aspects of the learning experience affects the extent of generalization, including the number of learning trials (Hearst, 1968; Hovland, 1937), reinforcement rate (Eliot Hearst, Koresko, & Poppen, 1964; Hull, 1947; Thomas & Switalski, 1966), or simple versus differential learning procedures (Baron, 1973; Chen, Barnes, & Wilson, 2011; Hanson, 1959; Jenkins & Harrison, 1962). For example, Hanson (1959) conducted one of the first controlled experiments of (intradimensional) differential learning. Pigeons were trained to peck for food during the presentation of a CS+ (reinforced conditioned stimulus), which was a hue with a wavelength of 550 nm. They were also presented with a CS- that was not paired with food (a hue with a different wavelength). The physical distance between the CS- and CS+ differed between groups. A control group was trained with the CS+ alone. When presented with a range of test stimuli (different hues), Hanson found marked differences between the groups in the shape of the obtained generalization gradient. Differential compared to simple learning (control group) led to steeper generalization gradients (i.e., stronger decrements in conditioned responding for stimuli that were more distinct from the CS+), with the steepness of the gradient increasing as the distance between the CS+ and CS- decreased.

Although the impact of those experimental variables has been well studied, it is far less understood (and acknowledged) how perceptual variability may influence learning and subsequently, generalization (but see Atkinson & Estes,

1963; Enquist & Ghirlanda, 2005). Despite being subjected to the same objective experimental procedure, not everyone will experience it in the same way for a variety of reasons, including subjects' past experiences, cognitive ability, attention, and perceptual accuracy. Imagine, as an example in extremis, a differential fear conditioning protocol with differently colored cues (green vs. red). The experience for a person with deuteranopia (i.e., a form of color blindness) will significantly diverge from that of a person with normal vision. The inability to discriminate colors will lead to the experience that a gray cue is followed 50% of the time by a shock, as opposed to an experience in which only the green cue was paired with a shock. With distinct cues (e.g., different geometrical forms) this issue is most likely not relevant. However, many studies use stimuli with small physical differences or bodily sensations as cues during learning, where perceptual mistakes may be frequent. A recent study demonstrated the implications of misperceptions on learning experiences (Zaman et al., 2017). In a differential fear conditioning paradigm, esophageal balloon distensions of different physical intensities (high and low) were used as CS+ and CS- (counterbalanced). Subjects had to categorize them upon stimulus presentation as either high or low based on perceived intensity. Unexpectedly, it was found that the high-intensity stimulus was misidentified as the low-intensity stimulus in approximately 50% of the trials. Yet, US reinforcement was administered irrespective of perception. In the group where the high-intensity stimulus served as the CS+, such misperception led to increased fear responding to both stimuli, perhaps because from the subjects' perspective, both stimuli were followed by the US. In the other group, where the high-intensity stimulus was used as the CS-, similar misidentifications were never followed by the US, and differential learning was observed, presumably because from the experience of the subject, only the low- but never the high-intensity stimulus was followed by the US. Analogously, Schroyen et al. (2015) used CSs that varied in intensity (i.e., resistive breathing loads) and found a non-differential increase in US expectancy in the group with the high-intensity stimulus as CS+. In this group, a flat generalization gradient was obtained located at the same height as the peak of the gradient of a successful differential fear learning group. Although no perceptual categorizations were recorded, their findings could be explained by potentially similar perceptual errors and their impact on the subjects' experience. These findings demonstrate that the congruent assessment of perception and behavior could enable a better understanding of the latter.

Individual variability among subjects in learning poses a challenge for generalization researchers because it impacts the shape of the generalization gradients. Mednick and Freedman (1960) wrote the following on this matter: "It is important that all studies attempting to demonstrate group differences in stimulus generalization responsiveness either equate their

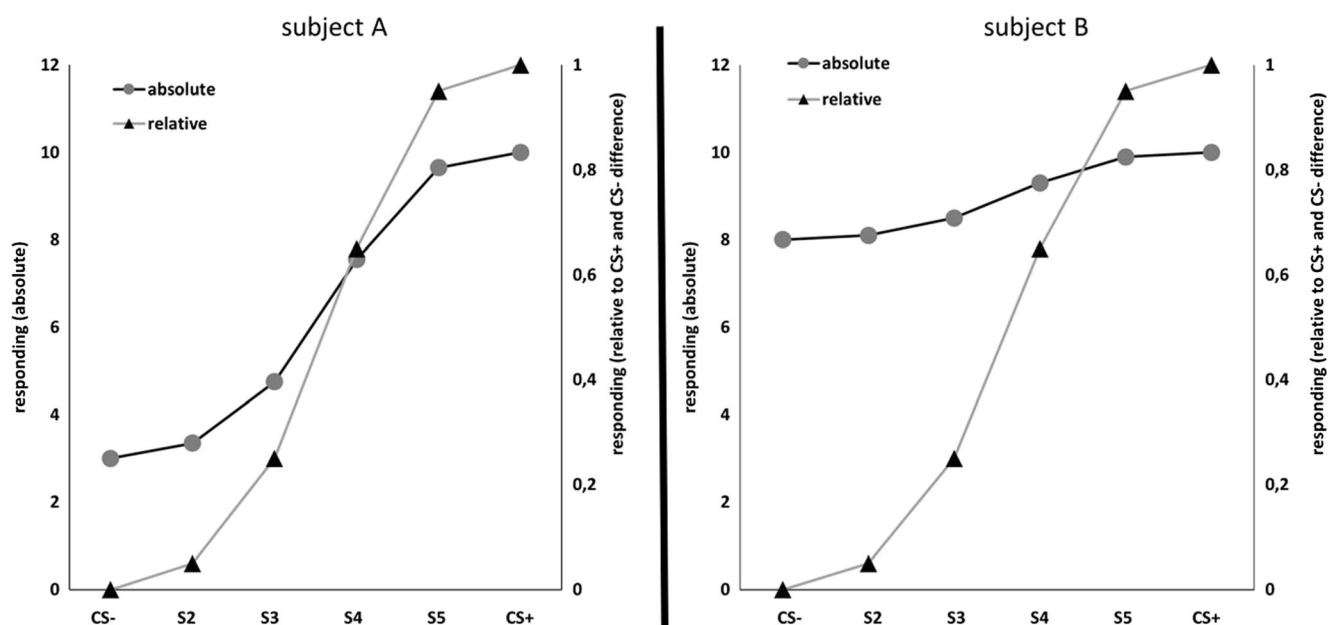


Fig. 8 Comparison of generalization gradients using absolute or relative conditioned responses (i.e., expressed in relation to the difference in response strength between the CS+ and CS-). CS conditioned stimulus, CS+ reinforced CS, CS- unreinforced CS, S test stimulus

groups for conditioning level or use a measure of relative generalization (ratio of generalization response to conditioned response)” (p. 185). To illustrate this, imagine two subjects (A and B) at the end of a differential fear learning phase: subject A had no problem to discriminate between stimuli perceptually, and learned that the CS+ was predictive of threat and the CS- predictive of its absence. Subject B, on the other hand, found it much more difficult to discriminate between the two stimuli and consequently, failed to pick up on these contingencies. Next, their responses to a range of stimuli located along the CS+/CS- continuum were recorded (Fig. 8). As per usual practice, response gradients were compared, without taking differences in learning into account, leading to the conclusion that subject B overgeneralizes compared to subject A. However, as both generalized from a different learning experience, inferring differences in generalization tendencies based on their absolute generalization gradient would be precarious. In Fig. 8, similar slopes between both are obtained when we express their responses in relation to the subject’s experience. A better interpretation of the data then would be that a failure to learn in subject B led to a (relatively) flat gradient (for an alternative interpretation, see Haddad, Pritchett, Lissek, & Lau, 2012). Subsequently, comparisons between subjects A and B’s absolute response gradients are not informative about interindividual differences regarding generalization tendencies.

This idea is not novel. It has been alluded to before (Mednick, 1957; Mednick & Freedman, 1960). Nevertheless, most contemporary generalization research compares, for instance, differences between patients and controls (Ahrens et al., 2016; Holt et al., 2014; Jensen et al., 2008;

Lissek et al., 2014; Morey et al., 2015; Reinecke, Becker, Hoyer, & Rinck, 2010; Tinoco-Gonzalez et al., 2015) or interindividual differences, regardless of possible differences in initial learning (Arnaudova, Kryptos, Effting, Kindt, & Beckers, 2017; Torrents-Rodas et al., 2013; Wills et al., 2011). Hence, reported differences in generalization should be interpreted cautiously, especially given, for example, reports of insufficient safety learning or differences in perceptual discrimination acuity in anxiety patients (Duits et al., 2015; Laufer et al., 2016). One approach to partly circumvent this issue is to make use of learning criteria (a response threshold that needs to be reached at the end of the training phase; Ginat-Frolich, Klein, Katz, & Shechner, 2017; Lommen, Engelhard, Sijbrandij, van den Hout, & Hermans, 2012). However, the use of such criteria can lead to the exclusion of (many) subjects (Lonsdorf et al., 2019), does not assure the absence of interindividual variation in learning, and can result in differential exclusion in, for example, patient groups versus control groups. Alternatively, a performance criterion could be combined with an adaptive learning phase, where the number of trials is dependent upon meeting certain criteria (Thomas, Mood, Morrison, & Wiertelak, 1991).

Box 2

As the perception of a stimulus is also a (mental) response, the question arises as to how these responses are related to conditioned responses. It may be that perceptual responses vary in the same way as a conditioned response, because they are both under control of the same third variable (e.g., training history). As an alternative, we propose a causal relationship between these two responses (i.e., a causal network of responses), in that perceptual responses are a determinant of conditioned responses. The findings that changes in perceptual discrimination were paralleled by neural alterations at primary sensory cortices

(Li et al., 2008), or that pharmacological inactivation of sensory cortices blocked conditioning-induced changes in discrimination acuity (Aizenberg & Geffen, 2013), suggest that changes in conditioning effects emerge due to altered sensory processing. However, the exact relation remains hypothetical as only a handful of studies measure both responses congruently.

It is our aim to inspire future research on this matter. As perception and generalization research have remained relatively separate fields, it remains unclear to which extent experimental manipulations known to affect generalization also influence perception and vice versa. To this end, the assessment of both perceptual and conditioned responses is crucial as inferences regarding either of them based on only one measure can be problematic. The systematic measurement of perceptual and conditioned responses would reveal the degree of covariance between these two responses, whether they are affected in the same manner by various experimental variables, and would enable researchers to identify the mechanisms via which experimental manipulations affect generalized responses. For instance, differences in generalization gradients due to differential rather than simple conditioning might actually be driven by changes in perception. However, without systematic research on this matter, it will remain unclear to what extent perceptual gradients and gradients in conditioned responding are driven by identical/different mechanisms, or which behavioral phenomena are actually observed because of perceptual variability.

Perceptual variability and memory

In the previous section, we discussed how perceptual variability could affect the learning experience. However, in order to affect future behavior, these experiences must be stored and retrieved from memory. In the following section, we briefly elaborate on how perceptual variability could affect memory and, in turn, how this could influence patterns of generalized responding.

A commonly held assumption is that poor memory results in broader generalization gradients (Jasnow, Cullen, & Riccio, 2012; Riccio, Ackil, & Burch-Vernon, 1992; Riccio, Richardson, & Ebner, 1984). This idea builds on the finding that longer delays between learning and testing flatten generalization gradients while preserving overall response strength (Desiderato, Butler, & Meyer, 1966; Perkins & Weyant, 1958; Riccio et al., 1984; Thomas & Burr, 1969). The forgetting of stimulus features due to the passage of time may increase the variability of a memory representation (Huang & Sekuler, 2010). Apart from forgetting, perceptual variability during learning may be another source that affects the width of the memory distribution. If the repeated presentation of a reinforced stimulus results in a wide range of percepts (or representations) on every trial, the result would be a broader distribution of units (or range of points in psychological space) that acquired a response potential (or associative strength). With similarity determined by the overlap in probability distributions (or point distances), a wide distribution will lead to more generalization compared to a narrow distribution (for an illustration, see Fig. 9) (Jasnow et al., 2012; Riccio et al.,

1992, 1984). In a recent study, we assessed perceptual memory of the CS using a recall task, before the assessment of generalization. We found that higher uncertainty ratings during memory recall led to broader US expectancy generalization gradients (Zenses et al., 2020 preprint).

As already mentioned, a common assumption in generalization research is that the representation of the trained stimulus is veridical. However, systematic perceptual variability may result in a shift of the distribution's mean from the actual stimulus, reflecting a memory bias, as the peak of the stimulus representation no longer corresponds to the actual stimulus. For example, biased memory could lead to a shift in response peak along the stimulus spectrum or *asymmetries* in the response gradient. Due to the shifted distribution of recalled CS features, other stimulus distributions share more overlap compared to the actual CS, which would result in increased responding (see Fig. 9). Indirect support for this comes from a study by Dunsmoor et al. (2009), where the test stimulus that elicited the strongest fear response was also most often identified as the trained stimulus during a recognition task (Dunsmoor, Mitroff, & LaBar, 2009). Furthermore, the perceptual categorization tasks in our studies also involve a memory component. As already mentioned, we found that for almost half of the subjects (58.6% of the sample in Zaman, Ceulemans, et al., 2019a and 47.5% of the sample in Zaman, Struyf, et al., 2019b) the peak of their perceptual gradients shifted away from the CS. In those studies, asymmetries in the gradient of conditioned responding were observed due to stronger conditioned responses in the direction of the potential memory bias (Struyf, Zaman, et al. 2017; in Supplementary Information of Zaman, Struyf, et al. 2019b). It remains to be determined whether these shifts are actually due to a memory bias. In the study discussed above where memory recall was assessed prior to generalization, we also found an effect of memory bias on the shape of the generalization gradient (Zenses et al., 2020 preprint).

Decision-making

Another determinant that has often been implicated – either explicitly or implicitly – in current generalization research relates to the notion that generalization comprises a decision-making process (Dunsmoor & Murphy, 2015; Dymond et al., 2015). Some associative theories indeed invoke separate performance mechanisms that only play a role at the time of testing (Boddez, Moors, Mertens, & De Houwer, 2020; Miller & Schachtman, 1988; Ralph R. Miller & Matzel, 1988). For example, it has been argued that conditioned responding is determined by a decision process that selects action tendencies that serve the goals of the organism (e.g., safety; for an elaborate discussion, see Boddez et al., 2020). To the extent that these goals are considered more important

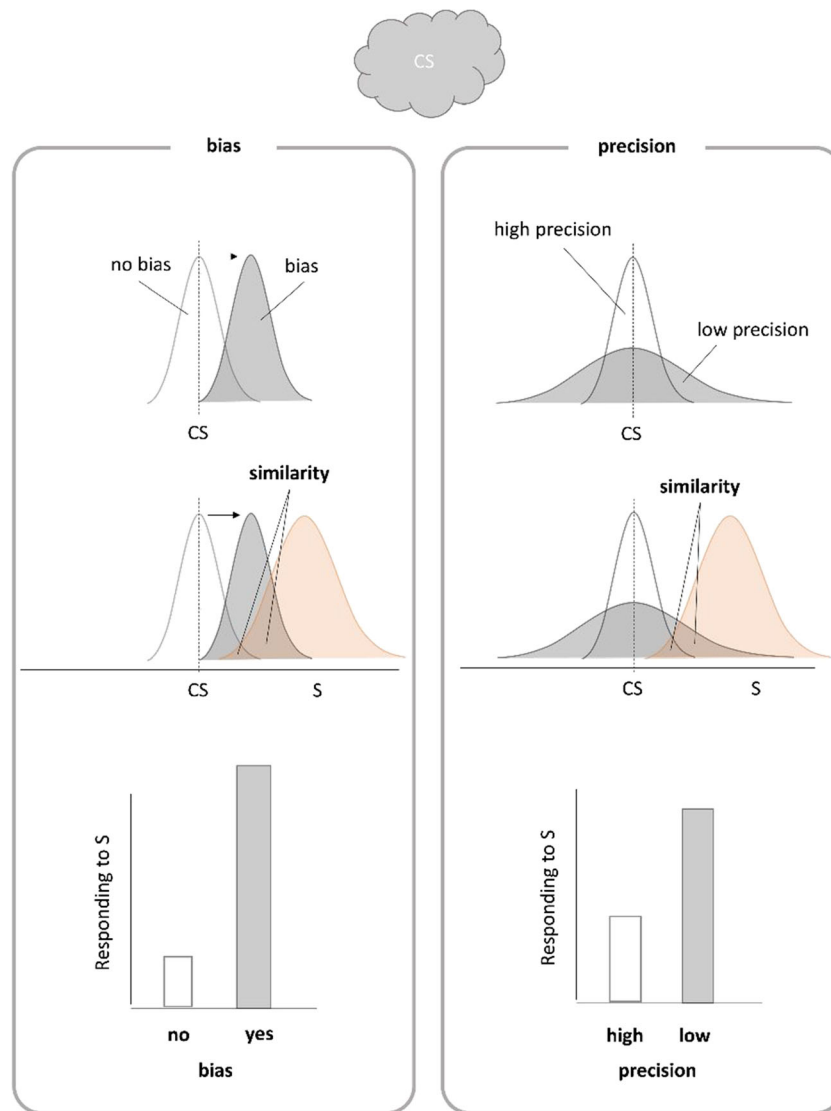


Fig. 9 Schematic example of the effects of precision and bias of the CS memory. **Left:** A biased CS memory results in more overlap between it and the representational distribution of a test stimulus (orange), yielding a higher level of similarity. The more similarity between the CS and a GS

the larger the elicited response. **Right:** A broad probability distributions of recalled CS features reflects a reduced memory precision. Thereby increasing its' overlap with the distribution of a test stimulus and, consequently, their similarity. *CS* conditioned stimulus, *S* test stimulus

(e.g., safety has a higher value) relative to other goals (e.g., saving energy), these action tendencies – and with them conditioned responding – will be set off more easily (e.g., in response to a GS). In addition, inferences concerning the situations in which these goals are either at stake or not will also affect responding. Differences in generalization gradients between specific patient populations and controls, for instance, have been attributed to a risk detection bias in the former (Dymond et al., 2015). Dymond et al. wrote the following on this matter: “This pattern of responding, referred to as overgeneralization, is defined by less steep decreases in fear responses as stimuli differentiate from CS+, indicating stronger generalization and the tendency to infer threat-potential to physically similar stimuli” (p. 563).

Furthermore, studies have demonstrated that different patterns of generalization are observed depending on instructions that may affect decision making (Ahmed & Lovibond, 2015; Vervliet, Kindt, Vansteenwegen, & Hermans, 2010). In the study of Vervliet et al. (2010), for example, two groups of subjects learned that a yellow triangle (CS+) was followed by shock (US). Crucially, instructions given prior to conditioning differed; in the Shape group (SG), the importance of shape in the occurrence of the US was emphasized, while in the Color group (CG), instructions suggested that color was essential for predicting the US. In a subsequent generalization test, a yellow square and a blue triangle were presented, while SCR and shock expectancies were measured. Findings revealed the vital significance of instructions, as those in the

SG showed more generalization to the blue triangle, while those in the CG generalized more to the yellow square. While both groups were exposed to the same stimulus pairings, verbal instructions influenced their generalized responding. In a follow-up study, Ahmed and Lovibond (2015) utilized the same paradigm, but this time, instructions were given after the conditioning phase to control for any possible effects on learning. Once again, similar results were obtained, confirming the impact of verbal instructions on generalization patterns.

Recent work found considerable heterogeneity among humans regarding adopted (and the ability to verbalize) rules. Different decision-rules were associated with distinct gradients in conditioned responding (Boddez, Bennett, van Esch, & Beckers, 2016; Lee, Lovibond, Hayes, & Navarro, 2018b; Lovibond et al., 2019). For instance, Lee et al. (2018a, b) found that subjects generalized using different decision strategies; they adopted either a similarity-based rule (i.e., “the more a stimulus resembles the CS+, the more dangerous it is”), a linear (or categorical) rule (i.e., “the bluer the stimulus is, the more dangerous it is”), or could not report any rule. The former decision strategy led to a gradient that peaked around the CS+, while the latter resulted in a linear gradient that peaked a bit further from the CS+, in the direction opposite from the CS-. Only through averaging of the gradients of these subgroups a peak shift in the overall gradient was found. Furthermore, explanations of the different response rules remain speculative, but we suggest that differences in perceptual accuracy may be one reason. For instance, when able to perceive the subtle differences between stimuli, the similarity-based rule matches the person’s experience. On the other hand, when only able to differentiate between the extreme stimuli, one may be more likely to develop a linear or categorical rule. After all, it only seems logical that the cognitive rules that we develop arise from our experiences when interacting with our environment rather than objective environmental settings.

Conclusion

The aim of this paper is to inspire new research by illustrating the various ways through which (non-static) perceptual processes can affect learning and generalization. In the literature, multiple aspects of generalization are studied but usually in isolation. More often than not, individual differences, among other things, are not taken into account, rendering the findings difficult to interpret. Future research should strive to attend to the sources of these individual differences, as well as to implement advances from other fields and novel analytical approaches, to help us improve the way we study generalization. More precisely, advances made in other fields (e.g., perception research) should not only be acknowledged but also implemented in current generalization models as they will benefit our understanding of underlying

processes involved in generalized responding. Taking perception and prior expectations into account in associative learning has already led to fruitful research in psychopathology. For instance, associative models of psychosis have argued that positive symptoms (e.g., delusions) might be due to perceptual disturbances caused by aberrant prediction error (Corlett et al., 2019; Powers, Mathys, & Corlett, 2017). In addition, we demonstrated that inferences regarding distinct gradients should be made cautiously as they could emerge via various ways. Finally, we advocated that future research should move away from population-based gradients and attempt to model and understand interindividual differences in gradients (Lee, Hayes, & Lovibond, 2018a; Struyf et al., 2015), as well as variations in responding to repeated stimulus presentations within the same subject. As such, analytical approaches other than just the analyses of averaged responses (per stimulus) might be better suited for generalization research (Vanbrabant, Raes, Hermans, & Vanpaemel, 2016), including the implementation of computational models that parametrize latent processes. This would aid researchers and clinicians to differentiate between various distinct processes underlying similar behavioral phenotypes, and as such, to further guide the development of individually tailored treatments.

Funding JZ is a postdoctoral Research Fellow of the Research Foundation Flanders (FWO, 12P8619N), and was supported by the “Asthenes” long-term structural funding (METH/15/011) - Methusalem grant by the Flemish Government, and a Krediet aan Navorsers (FWO, 1500620N). YB is employed on a Methusalem Grant (BOF16/MET_V/002).

References

- Ahmed, O., & Lovibond, P. F. (2015). The Impact of Instructions on Generalization of Conditioned Fear in Humans. *Behavior Therapy*, 46(5), 597–603. <https://doi.org/10.1016/j.beth.2014.12.007>
- Ahrens, L. M., Pauli, P., Reif, A., Mühlberger, AndreasLangs, G., Aalderink, T., & Wieser, M. J. (2016). Fear conditioning and stimulus generalization in patients with social anxiety disorder. *Journal of Anxiety Disorders*, 44, 36–46. <https://doi.org/10.1016/j.janxdis.2016.10.003>
- Åhs, F., Miller, S. S., Gordon, A. R., & Lundström, J. N. (2013). Aversive learning increases sensory detection sensitivity. *Biological Psychology*, 92(2), 135–141. <https://doi.org/10.1016/j.biopsycho.2012.11.004>
- Aizenberg, M., & Geffen, M. N. (2013). Bidirectional effects of aversive learning on perceptual acuity are mediated by the sensory cortex. *Nature Neuroscience*, 16(8), 994–996. <https://doi.org/10.1038/nn.3443>
- Ames, L. L., & Yarczower, M. (1965). Some effects of wavelength discrimination on stimulus generalization in the goldfish. *Psychonomic Science*, 3(1–12), 311–312.
- Armony, J. L., Servan-Schreiber, D., Romanski, L. M., Cohen, J. D., & LeDoux, J. E. (1997). Stimulus generalization of fear responses: effects of auditory cortex lesions in a computational model and in rats. *Cerebral Cortex (New York, N.Y. : 1991)*, 7(2), 157–165. Retrieved from <http://www.ncbi.nlm.nih.gov/pubmed/9087823>
- Amadova, I., Kryptos, A. M., Effting, M., Kindt, M., & Beckers, T. (2017). Fearing shades of grey: individual differences in fear

- responding towards generalisation stimuli. *Cognition and Emotion*, 31(6), 1181–1196. <https://doi.org/10.1080/02699931.2016.1204990>
- Ashby, F. G., & Lee, W. W. (1993). Perceptual Variability as a Fundamental Axiom of Perceptual Science (pp. 369–399). [https://doi.org/10.1016/S0166-4115\(08\)62778-8](https://doi.org/10.1016/S0166-4115(08)62778-8)
- Ashby, F. G., & Maddox, W. T. (1993). Relations between Prototype, Exemplar, and Decision Bound Models of Categorization. *Journal of Mathematical Psychology*, 37(3), 372–400. <https://doi.org/10.1006/jmps.1993.1023>
- Ashby, F. G., & Townsend, J. T. (1986). Varieties of perceptual independence. *Psychological Review*, 93(2), 154–179. <https://doi.org/10.1037/0033-295X.93.2.154>
- Asutay, E., & Västfjäll, D. (2012). Perception of loudness is influenced by emotion. *PLoS ONE*, 7(6), 2–6. <https://doi.org/10.1371/journal.pone.0038660>
- Atkinson, R. C., & Estes, W. K. (1963). Stimulus sampling theory. In R. D. Luce, R. R. Bush, & E. Galanter (Eds.), *Handbook of mathematical psychology* (3rd, pp. 121–268). New York: Wiley.
- Baron, A. (1973). Postdiscrimination gradients of human subjects on a tone continuum. *Journal of Experimental Psychology*, 101(2), 337–342. <https://doi.org/10.1037/h0035206>
- 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), 3–21.
- Blough, P. M. (1972). Wavelength generalization and discrimination in the pigeon. *Perception & Psychophysics*, 12(4), 342–348. <https://doi.org/10.3758/BF03207219>
- Boddez, Y., Moors, A., Mertens, G., & De Houwer, J. (2020). Tackling fear: Beyond associative memory activation as the only determinant of fear responding. *Neuroscience & Biobehavioral Reviews*, 112, 4210–4419.
- Boddez, Yannick, Bennett, M. P., van Esch, S., & Beckers, T. (2016). Bending rules: the shape of the perceptual generalisation gradient is sensitive to inference rules. *Cognition and Emotion*, 1–9. <https://doi.org/10.1080/02699931.2016.1230541>
- Chambers, C., & Pressnitzer, D. (2014). Perceptual hysteresis in the judgment of auditory pitch shift. *Attention, Perception, & Psychophysics*, 76(5), 1271–1279. <https://doi.org/10.3758/s13414-014-0676-5>
- Chen, C.-F. F., Barnes, D. C., & Wilson, D. A. (2011). Generalized vs. stimulus-specific learned fear differentially modifies stimulus encoding in primary sensory cortex of awake rats. *Journal of Neurophysiology*, 106(6), 3136–3144. <https://doi.org/10.1152/jn.00721.2011>
- Cicchini, G. M., Mikellidou, K., & Burr, D. (2017). Serial dependencies act directly on perception. *Journal of Vision*, 17(14), 6. <https://doi.org/10.1167/17.14.6>
- Clark, A. (2013). Whatever next? Predictive brains, situated agents, and the future of cognitive science. *Behavioral and Brain Sciences*, 36(3), 181–204. <https://doi.org/10.1017/S0140525X12000477>
- Cohen, J. D., LeDoux, J. E., Armony, J. L., & Servan-Schreiber, D. (1997). Computational modeling of emotion: explorations through the anatomy and physiology of fear conditioning. *Trends in Cognitive Sciences*, 1(1), 28–34. [https://doi.org/10.1016/S1364-6613\(97\)01007-3](https://doi.org/10.1016/S1364-6613(97)01007-3)
- Corlett, P. R., Horga, G., Fletcher, P. C., Alderson-Day, B., Schmack, K., & Powers, A. R. (2019). Hallucinations and Strong Priors. *Trends in Cognitive Sciences*, 23(2), 114–127. <https://doi.org/10.1016/j.tics.2018.12.001>
- Davis, T., & Love, B. C. (2010). Memory for Category Information Is Idealized Through Contrast With Competing Options. *Psychological Science*, 21(2), 234–242. <https://doi.org/10.1177/0956797609357712>
- de Houwer, J., Barnes-Holmes, D., & Moors, A. (2013). What is learning? On the nature and merits of a functional definition of learning. *Psychonomic Bulletin and Review*, 20(4), 631–642. <https://doi.org/10.3758/s13423-013-0386-3>
- De Hoz, L., & Nelken, I. (2014). Frequency tuning in the behaving mouse: Different bandwidths for discrimination and generalization. *PLoS ONE*, 9(3). <https://doi.org/10.1371/journal.pone.0091676>
- Desiderato, O., Butler, B., & Meyer, C. (1966). Changes in fear generalization gradients as a function of delayed testing. *Journal of Experimental Psychology*, 72(5), 678–682. <https://doi.org/10.1037/h0023798>
- Diesch, E., & Flor, H. (2007). Alteration in the response properties of primary somatosensory cortex related to differential aversive Pavlovian conditioning. *Pain*, 131(1–2), 171–180. <https://doi.org/10.1016/j.pain.2007.01.016>
- Dong, X., Gao, Y., Lv, L., & Bao, M. (2016). Habituation of visual adaptation. *Scientific Reports*, 6(1), 19152. <https://doi.org/10.1038/srep19152>
- Duits, P., Cath, D. C., Lissek, S., Hox, J. J., Hamm, A. O., Engelhard, I. M., ... Baas, J. M. P. (2015). Updated meta-analysis of classical fear conditioning in the anxiety disorders. *Depression and Anxiety*, 32(4), 239–253. <https://doi.org/10.1002/da.22353>
- Dunsmoor, J. E., Mitroff, S. R., & LaBar, K. S. (2009). Generalization of conditioned fear along a dimension of increasing fear intensity. *Learning and Memory*, 16(7), 460–469. <https://doi.org/10.1101/lm.1431609>
- Dunsmoor, J. E., & Murphy, G. L. (2015). Categories, concepts, and conditioning: how humans generalize fear. *Trends in Cognitive Sciences*, 19(2), 73–77. <https://doi.org/10.1016/j.tics.2014.12.003>
- Dunsmoor, J. E., White, A. J., & LaBar, K. S. (2011). Conceptual similarity promotes generalization of higher order fear learning. *Learning & Memory (Cold Spring Harbor, N.Y.)*, 18(3), 156–160. <https://doi.org/10.1101/lm.2016411>
- Dymond, S., Dunsmoor, J. E., Vervliet, B., Roche, B., & Hermans, D. (2015). Fear Generalization in Humans: Systematic Review and Implications for Anxiety Disorder Research. *Behavior Therapy*, 46(5), 561–582. <https://doi.org/10.1016/j.beth.2014.10.001>
- Enquist, M., & Ghirlanda, S. (2005). Neural Networks and Animal Behavior (Vol. 33). Princeton: Princeton University Press. <https://doi.org/10.1515/9781400850785>
- Estes, W. K., & Burke, C. J. (1953). A theory of stimulus variability in learning. *Psychological Review*, 60(4), 276–286. <https://doi.org/10.1037/h0055775>
- Estes, W. K. (1956). The problem of inference from curves based on group data. *Psychological Bulletin*, 53(2), 134–140. <https://doi.org/10.1037/h0045156>
- Feldman, N. H., Griffiths, T. L., & Morgan, J. L. (2009). Perceptual Magnet Effect As Optimal Statistical Inference. *Psychological Review*, 116(4), 752–782. <https://doi.org/10.1037/a0017196>
- Friston, K., & Kiebel, S. (2009). Predictive coding under the free-energy principle. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 364(1521), 1211–1221. <https://doi.org/10.1098/rstb.2008.0300>
- Ganz, L., & Riesen, A. H. (1962). Stimulus generalization to hue in the darkreared macaque. *Journal of Comparative and Physiological Psychology*, 55(1), 92–99. <https://doi.org/10.1037/h0044987>
- Ghirlanda, S. (2002). Intensity generalization: Physiology and modelling of a neglected topic. *Journal of Theoretical Biology*, 214(3), 389–404. <https://doi.org/10.1006/jtbi.2001.2452>
- Ghirlanda, S. (2015). On elemental and configural models of associative learning. *Journal of Mathematical Psychology*, 64–65, 8–16. <https://doi.org/10.1016/j.jmp.2014.11.003>
- Ghirlanda, S., & Enquist, M. (2003). A century of generalization. *Animal Behaviour*, 66(1), 15–36. <https://doi.org/10.1006/anbe.2003.2174>
- Ghirlanda, S., & Enquist, M. (2007). How training and testing histories affect generalization: A test of simple neural networks. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 362(1479), 449–454. <https://doi.org/10.1098/rstb.2006.1972>
- Gibson, E. (1969). *Principles of perceptual learning and development*. Appleton-Century-Crofts.

- Gilbert, C. D., Sigman, M., & Crist, R. E. (2001). The neural basis of perceptual learning. *Neuron*, 31(5), 681–697. [https://doi.org/10.1016/S0896-6273\(01\)00424-X](https://doi.org/10.1016/S0896-6273(01)00424-X)
- Ginat-Frolich, R., Klein, Z., Katz, O., & Shechner, T. (2017). A novel perceptual discrimination training task: Reducing fear overgeneralization in the context of fear learning. *Behaviour Research and Therapy*, 93, 29–37. <https://doi.org/10.1016/j.brat.2017.03.010>
- Guttman, N., & Kalish, H. I. (1956). Discriminability and stimulus generalization. *Journal of Experimental Psychology*, 51(1), 79–88. <https://doi.org/10.1037/h0046219>
- Haddad, A. D. M., Pritchett, D., Lissek, S., & Lau, J. Y. F. (2012). Trait anxiety and fear responses to safety cues: Stimulus generalization or sensitization? *Journal of Psychopathology and Behavioral Assessment*, 34(3), 323–331. <https://doi.org/10.1007/s10862-012-9284-7>
- Hamilton, W. F., & Coleman, T. B. (1933). Trichromatic vision in the pigeon as illustrated by the spectral hue discrimination curve. *Journal of Comparative Psychology*, 15(1), 183–191. <https://doi.org/10.1037/h0073166>
- Hanson, H. M. (1957). Discrimination Training Effect on Stimulus Generalization Gradient for Spectrum Stimuli. *Science*, 125(3253), 888–889. <https://doi.org/10.1126/science.125.3253.888>
- Hanson, Harley M. (1959). Effects of discrimination training on stimulus generalization. *Journal of Experimental Psychology*, 58(5), 321–334. <https://doi.org/10.1037/h0042606>
- Hayes, K. J. (1953). The backward curve: a method for the study of learning. *Psychological Review*, 60(4), 269–275. <https://doi.org/10.1037/h0056308>
- Hearst, E. (1968). Discrimination Learning as the Summation of Excitation and Inhibition. *Science*, 162(3859), 1303–1306. <https://doi.org/10.1126/science.162.3859.1303>
- Hearst, Eliot, Koresko, M. B., & Poppen, R. (1964). Stimulus generalization and the response-reinforcement contingency. *Journal of the Experimental Analysis of Behavior*, 7(5), 369–380.
- Holt, D. J., Boeke, E. A., Wolthuisen, R. P. F., Nasr, S., Milad, M. R., & Tootell, R. B. H. (2014). A parametric study of fear generalization to faces and non-face objects: relationship to discrimination thresholds. *Frontiers in Human Neuroscience*, 8, 1–12. <https://doi.org/10.3389/fnhum.2014.00624>
- Honig, W. K., & Urciuoli, P. J. (1981). The legacy of Guttman and Kalish (1956): Twenty-five years of research on stimulus generalization. *Journal of the Experimental Analysis of Behavior*, 36(3), 405–445. <https://doi.org/10.1901/jeab.1981.36-405>
- Hoskin, R., Berzuini, C., Acosta-Kane, D., El-Derey, W., Guo, H., & Talmi, D. (2019). Sensitivity to pain expectations: A Bayesian model of individual differences. *Cognition*, 182(November 2017), 127–139. <https://doi.org/10.1016/j.cognition.2018.08.022>
- Hovland, C. I. (1937). The Generalization of Conditioned Responses: II. The Sensory Generalization of Conditioned Responses with Varying Intensities of Tone. *Pedagogical Seminary and Journal of Genetic Psychology*, 51(2), 279–291. <https://doi.org/10.1080/08856559.1937.10532503>
- Huang, J., & Sekuler, R. (2010). Distortions in recall from visual memory: two classes of attractors at work. *Journal of Vision*, 10(2), 24.1–27. <https://doi.org/10.1167/10.2.24>
- Hull, C. L. (1947). The problem of primary stimulus generalization. *Psychological Review*, 54(3), 120–134. <https://doi.org/10.1037/h0061159>
- Jäkel, F., Schölkopf, B., & Wichmann, F. A. (2008). Generalization and similarity in exemplar models of categorization: Insights from machine learning. *Psychonomic Bulletin and Review*, 15(2), 256–271. <https://doi.org/10.3758/PBR.15.2.256>
- Jasnow, A. M., Cullen, P. K., & Riccio, D. C. (2012). Remembering another aspect of forgetting. *Frontiers in Psychology*, 3(JUN), 1–8. <https://doi.org/10.3389/fpsyg.2012.00175>
- Jenkins, H. M., & Harrison, R. H. (1962). Generalization gradients of inhibition following auditory discrimination learning. *Journal of the Experimental Analysis of Behavior*, 5(4), 435–441.
- Jensen, J., Willeit, M., Zipursky, R. B., Savina, I., Smith, A. J., Menon, M., ... Kapur, S. (2008). The formation of abnormal associations in schizophrenia: Neural and behavioral evidence. *Neuropsychopharmacology*, 33(3), 473–479. <https://doi.org/10.1038/sj.npp.1301437>
- Jones, M., Love, B. C., & Maddox, W. T. (2006). Recency effects as a window to generalization: Separating decisional and perceptual sequential effects in category learning. *Journal of Experimental Psychology: Learning Memory and Cognition*, 32(2), 316–332. <https://doi.org/10.1037/0278-7393.32.3.316>
- Jones, M., & Sieck, W. R. (2003). Learning myopia: An adaptive recency effect in category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 29(4), 626–640. <https://doi.org/10.1037/0278-7393.29.4.626>
- Kalish, H. I. (1958). The relationship between discriminability and generalization: A re-evaluation. *Journal of Experimental Psychology*, 55(6), 637–644. <https://doi.org/10.1037/h0048049>
- Kass, M. D., Rosenthal, M. C., Pottackal, J., & McGann, J. P. (2013). Fear learning enhances neural responses to threat-predictive sensory stimuli. *Science (New York, N.Y.)*, 342(6164), 1389–1392. <https://doi.org/10.1126/science.1244916>
- Kowal, K. H. (1993). The range effect as a function of stimulus set, presence of a standard, and modulus. *Perception & Psychophysics*, 54(4), 555–561. <https://doi.org/10.3758/BF03211777>
- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99(1), 22–44. <https://doi.org/10.1037/0033-295X.99.1.22>
- Lange, I., Goossens, L., Michiels, S., Bakker, J., Lissek, S., Papalini, S., ... Schruers, K. (2017). Behavioral pattern separation and its link to the neural mechanisms of fear generalization. *Social Cognitive and Affective Neuroscience*, 12(11), 1720–1729. <https://doi.org/10.1093/scan/nsx104>
- Lashley, K. S., & Wade, M. (1946). The Pavlovian theory of generalization. *Psychological Review*, 53(2), 72–87. <https://doi.org/10.1037/h0059999>
- Laufer, O., Israeli, D., & Paz, R. (2016). Behavioral and neural mechanisms of overgeneralization in anxiety. *Current Biology*, 26(6), 713–722. <https://doi.org/10.1016/j.cub.2016.01.023>
- Laufer, O., & Paz, R. (2012). Monetary loss alters perceptual thresholds and compromises future decisions via amygdala and prefrontal networks. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(18), 6304–6311. <https://doi.org/10.1523/JNEUROSCI.6281-11.2012>
- Lee, J. C., Hayes, B. K., & Lovibond, P. F. (2018a). Peak Shift and Rules in Human Generalization. *Journal of Experimental Psychology: Learning Memory and Cognition*, (January). <https://doi.org/10.1037/xlm0000558>
- Lee, J. C., Lovibond, P. F., Hayes, B. K., & Navarro, D. J. (2018b). Negative Evidence and Inductive Reasoning in Generalization of Associative Learning The role of inductive reasoning in generalization of human associative learning View project. *Article in Journal of Experimental Psychology General*, 148(2), 289–303. <https://doi.org/10.1037/xge0000496>
- Li, W., Howard, J. D., Parrish, T. B., & Gottfried, J. A. (2008). Aversive learning enhances perceptual and cortical discrimination of indiscriminable odor cues. *Science*, 319(5871), 1842–1845. <https://doi.org/10.1126/science.1152837>
- 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(5), 678–687. <https://doi.org/10.1016/j.brat.2008.02.005>

- Lissek, S., & Grillon, C. (2010). Overgeneralization of Conditioned Fear in the Anxiety Disorders. *Zeitschrift Für Psychologie / Journal of Psychology*, 218(2), 146–148. <https://doi.org/10.1027/0044-3409/a000022>
- Lissek, S., Kaczurkin, A. N., Rabin, S., Geraci, M., Pine, D. S., & Grillon, C. (2014). Generalized anxiety disorder is associated with overgeneralization of classically conditioned-fear. *Biological Psychiatry*, 75(11), 909–915. <https://doi.org/10.1016/j.biopsych.2013.07.025>
- Lissek, S., Rabin, S., Heller, R. E., Lukenbaugh, D., Geraci, M., Pine, D. S., & Grillon, C. (2010). Overgeneralization of conditioned fear as a pathogenic marker of panic disorder. *The American Journal of Psychiatry*, 167(1), 47–55. <https://doi.org/10.1176/appi.ajp.2009.09030410>
- Lommen, M. J. J., Engelhard, I. M., Sijbrandij, M., van den Hout, M. A., & Hermans, D. (2012). Pre-trauma individual differences in extinction learning predict posttraumatic stress. *Behaviour Research and Therapy*, 51(2), 63–67. <https://doi.org/10.1016/j.brat.2012.11.004>
- Lonsdorf, T. B., Klingelhöfer-Jens, M., Andreatta, M., Beckers, T., Chalkia, A., Gerlicher, A., ... Merz, C. J. (2019). Navigating the garden of forking paths for data exclusions in fear conditioning research. *eLife*, 8. <https://doi.org/10.7554/eLife.52465>
- Lovibond, P. F., Lee, J. C., & Hayes, B. K. (2019). Stimulus discriminability and induction as independent components of generalization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, (September). <https://doi.org/10.1037/xlm0000779>
- Maddox, W. T. (1999). On the dangers of averaging across observers when comparing decision bound models and generalized context models of categorization. *Perception and Psychophysics*, 61(2), 354–374. <https://doi.org/10.3758/BF03206893>
- McLaren, I. P. L., & Mackintosh, N. J. (2000). An elemental model of associative learning: I. Latent inhibition and perceptual learning. *Animal Learning & Behavior*, 28(3), 211–246. <https://doi.org/10.3758/BF03200258>
- McLaren, I. P. L., & Mackintosh, N. J. (2002). Associative learning and elemental representation: II. Generalization and discrimination. *Animal Learning & Behavior*, 30(3), 177–200. <https://doi.org/10.3758/BF03192828>
- Mednick, S. A. (1957). Generalization as a function of manifest anxiety and adaptation to psychological experiments. *Journal of Consulting Psychology*, 21(6), 491–494. <https://doi.org/10.1037/h0045843>
- Mednick, S. A., & Freedman, J. L. (1960). Stimulus generalization. *Psychological Bulletin*, 57(3), 169–200. <https://doi.org/10.1037/h0041650>
- Merrell M. (1931) The relationship of individual growth to average growth. *Human Biology* 3, 37–70
- Miller, R. R., & Schachtman, T. R. (1988). Conditioning context as an associative baseline: implications for response generation and the nature of conditioned inhibition. In G. H. Bower (Ed.), *The Psychology of Learning and Motivation* (22nd, pp. 51–92). San Diego: Academic Press.
- Miller, Ralph R., & Matzel, L. D. (1988). The Comparator Hypothesis: A Response Rule for The Expression of Associations (pp. 51–92). 10.1016/S0079-7421(08)60038-9
- Morey, R. A., Dunsmoor, J. E., Haswell, C. C., Brown, V. M., Vora, A., Weiner, J., ... Iii, H. R. W. (2015). Fear learning circuitry is biased toward generalization of fear associations in posttraumatic stress disorder. 5(12), e700-10. <https://doi.org/10.1038/tp.2015.196>
- Muensinger, J., Stingl, K. T., Matuz, T., Binder, G., Ehehalt, S., & Preissl, H. (2013). Auditory habituation to simple tones: Reduced evidence for habituation in children compared to adults. *Frontiers in Human Neuroscience*, 7(JUL), 1–7. <https://doi.org/10.3389/fnhum.2013.00377>
- Nosofsky, R. M. (1986). Attention, Similarity, and the Identification-Categorization Relationship. *Journal of Experimental Psychology: General*, 115(1), 39–57. <https://doi.org/10.1037/0096-3445.115.1.39>
- Onat, S., & Büchel, C. (2015). The neuronal basis of fear generalization in humans. *Nature Neuroscience*, 18(12), 1811–1818. <https://doi.org/10.1038/nn.4166>
- Ons, B., De Baene, W., & Wagemans, J. (2011). Subjectively Interpreted Shape Dimensions as Privileged and Orthogonal Axes in Mental Shape Space. *Journal of Experimental Psychology: Human Perception and Performance*, 37(2), 422–441. <https://doi.org/10.1037/a0020405>
- Pérez-González, D., & Malmierca, M. S. (2014). Adaptation in the auditory system: An overview. *Frontiers in Integrative Neuroscience*, 8(FEB), 1–10. <https://doi.org/10.3389/fnint.2014.00019>
- Perkins, C. C., & Weyant, R. G. (1958). The interval between training and test trials as a determiner of the slope of generalization gradients. *Journal of Comparative and Physiological Psychology*, 51(5), 596–600. <https://doi.org/10.1037/h0042550>
- Petzschner, F. H., & Glasauer, S. (2011). Iterative Bayesian estimation as an explanation for range and regression effects: A study on human path integration. *Journal of Neuroscience*, 31(47), 17220–17229. <https://doi.org/10.1523/JNEUROSCI.2028-11.2011>
- Petzschner, F. H., Glasauer, S., & Stephan, K. E. (2015). A Bayesian perspective on magnitude estimation. *Trends in Cognitive Sciences*, 19(5), 285–293. <https://doi.org/10.1016/j.tics.2015.03.002>
- Powers, A. R., Mathys, C., & Corlett, P. R. (2017). Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science*, 357(6351), 596–600. <https://doi.org/10.1126/science.aan3458>
- Press, C., Kok, P., & Yon, D. (2020). The Perceptual Prediction Paradox. *Trends in Cognitive Sciences*, 24(1), 13–24. <https://doi.org/10.1016/j.tics.2019.11.003>
- Purtle, R. B. (1973). Peak shift: A review. *Psychological Bulletin*, 80(5), 408–421. <https://doi.org/10.1037/h0035233>
- Reinecke, A., Becker, E. S., Hoyer, J., & Rinck, M. (2010). Generalized implicit fear associations in generalized anxiety disorder. *Depression and Anxiety*, 27(3), 252–259. <https://doi.org/10.1002/da.20662>
- Resnik, J., Sobel, N., & Paz, R. (2011). Auditory aversive learning increases discrimination thresholds. *Nature Neuroscience*, 14(6), 791–796. <https://doi.org/10.1038/nn.2802>
- Riccio, D. C., Ackil, J., & Burch-Vernon, A. (1992). Forgetting of stimulus attributes: methodological implications for assessing associative phenomena. *Psychological Bulletin*, 112(3), 433–445. <https://doi.org/10.1037/0033-2909.112.3.433>
- Riccio, D. C., Richardson, R., & Ebner, D. L. (1984). Memory retrieval deficits based upon altered contextual cues: A paradox. *Psychological Bulletin*, 96(1), 152–165. <https://doi.org/10.1037/0033-2909.96.1.152>
- Sagi, D. (2011). Perceptual learning in Vision Research. *Vision Research*, 51(13), 1552–1566. <https://doi.org/10.1016/j.visres.2010.10.019>
- Samaey, C., Wagemans, J., & Moors, P. (2020). Individual differences in processing orientation and proximity as emergent features. *Vision Research*, 169, 12–24. <https://doi.org/10.1016/j.visres.2020.02.002>
- Schechtman, E., Laufer, O., & Paz, R. (2010). Negative valence widens generalization of learning. *The Journal of Neuroscience : The Official Journal of the Society for Neuroscience*, 30(31), 10460–10464. <https://doi.org/10.1523/JNEUROSCI.2377-10.2010>
- Schroijen, M., Pappens, M., Schruers, K., Van den Bergh, O., Vervliet, B., & Van Diest, I. (2015). Generalization of Fear to Respiratory Sensations. *Behavior Therapy*, 46(5), 611–626. <https://doi.org/10.1016/j.beth.2015.05.004>
- Shepard, R. (1987). Toward a universal law of generalization for psychological science. *Science*, 237(4820), 1317–1323. <https://doi.org/10.1126/science.3629243>

- Shepard, R. N. (1958). Stimulus and response generalization: Deduction of the generalization gradient from a trace model. *Psychological Review*, 65(4), 242–256. <https://doi.org/10.1037/h0043083>
- Shepard, R. N., & Chang, J. J. (1963). Stimulus generalization in the learning of classifications. *Journal of Experimental Psychology*, 65(1), 94–102. <https://doi.org/10.1037/h0043732>
- Sidman, M. (1952). A note on functional relations obtained from group data. *Psychological Bulletin*, 49(3), 263–269. <https://doi.org/10.1037/h0063643>
- Solomon, P. R., & Moore, J. W. (1975). Latent inhibition and stimulus generalization of the classically conditioned nictitating membrane response in rabbits (*Oryctolagus cuniculus*) following hippocampal ablation. *Journal of Comparative and Physiological Psychology*, 89(10), 1192–1203. <https://doi.org/10.1037/h0077183>
- Spence, K. W. (1937). The differential response in animals to stimuli varying within a single dimension. *Psychological Review*, 44(5), 430–444. <https://doi.org/10.1037/h0062885>
- Stolarova, M., Keil, A., & Moratti, S. (2006). Modulation of the C1 visual event-related component by conditioned stimuli: evidence for sensory plasticity in early affective perception. *Cerebral Cortex (New York, N.Y. : 1991)*, 16(6), 876–887. <https://doi.org/10.1093/cercor/bhj031>
- Struyf, D., Zaman, J., Hermans, D., & Vervliet, B. (2017). Gradients of fear: How perception influences fear generalization. *Behaviour Research and Therapy*, 93, 116–122. <https://doi.org/10.1016/j.brat.2017.04.001>
- Struyf, D., Zaman, J., Vervliet, B., & Van Diest, I. (2015). Perceptual discrimination in fear generalization: Mechanistic and clinical implications. *Neuroscience and Biobehavioral Reviews*, 59, 201–207. <https://doi.org/10.1016/j.neubiorev.2015.11.004>
- Tenenbaum, J. B., & Griffiths, T. L. (2001). Generalization, similarity, and Bayesian inference. *Behavioral and Brain Sciences*, 24(4), 629–640. <https://doi.org/10.1017/S0140525X01000061>
- Thomas, D. R., & Burr, D. E. S. (1969). Stimulus generalization as a function of the delay between training and testing procedures: A reevaluation. *Journal of the Experimental Analysis of Behavior*, 12(1), 105–109. <https://doi.org/10.1901/jeab.1969.12-105>
- 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), 130–140. <https://doi.org/10.1037/0097-7403.17.2.130>
- Thomas, D. R., & Switalski, R. W. (1966). Comparison of stimulus generalization following variable-ratio and variable-interval training. *Journal of Experimental Psychology*, 71(2), 236–240. <https://doi.org/10.1037/h0022880>
- Tinoco-Gonzalez, D., Fullana, M. A., Torrents-Rodas, D., Bonillo, A., Vervliet, B., Blasco, M. J., ... Torrubia, R. (2015). Conditioned Fear Acquisition and Generalization in Generalized Anxiety Disorder. *Behavior Therapy*, 46(5), 627–639. <https://doi.org/10.1016/j.beth.2014.12.004>
- Torrents-Rodas, D., Fullana, M. A., Bonillo, A., Caseras, X., Andión, O., & Torrubia, R. (2013). No effect of trait anxiety on differential fear conditioning or fear generalization. *Biological Psychology*, 92(2), 185–190. <https://doi.org/10.1016/j.biopsycho.2012.10.006>
- Vanbrabant, K., Raes, F., Hermans, D., & Vanpaemel, W. (2016). *Hierarchical Models for Generalization Gradients: A Bayesian Approach*. KU Leuven.
- Vervliet, B., Iberico, C., Vervoort, E., & Baeyens, F. (2011). Generalization gradients in human predictive learning: Effects of discrimination training and within-subjects testing. *Learning and Motivation*, 42(3), 210–220. <https://doi.org/10.1016/j.lmot.2011.03.004>
- Vervliet, B., Kindt, M., Vansteenwegen, D., & Hermans, D. (2010). Fear generalization in humans: Impact of prior non-fearful experiences. *Behaviour Research and Therapy*, 48(11), 1078–1084. <https://doi.org/10.1016/j.brat.2010.07.002>
- Wiech, K., Vandekerckhove, J., Zaman, J., Tuerlinckx, F., Vlaeyen, J. W. S., J. W. S. J. W. S., Tracey, I., ... Tracey, I. (2014). Influence of prior information on pain involves biased perceptual decision-making. *Current Biology*, 24(15), R679–R681. <https://doi.org/10.1016/j.cub.2014.06.022>
- Wills, A. J., Road, P., Uk, E. E. X., Road, P., Uk, E. E. X., & Iplmclarenexacuk, I. P. L. M. (2011). Proceedings of the Annual Meeting of the Cognitive Science Working Memory Capacity and Generalization in Predictive Learning.
- Zaman, J., Ceulemans, E., Hermans, D., & Beckers, T. (2019a). Direct and indirect effects of perception on generalization gradients. *Behaviour Research and Therapy*, 114(January), 44–50. <https://doi.org/10.1016/j.brat.2019.01.006>
- Zaman, J., Struyf, D., Ceulemans, E., Beckers, T., & Vervliet, B. (2019b). Probing the role of perception in fear generalization. *Scientific Reports*, 9(1), 10026. <https://doi.org/10.1038/s41598-019-46176-x>
- Zaman, J., Struyf, D., Ceulemans, E., Vervliet, B., & Beckers, T. (2020). Perceptual errors are related to shifts in generalization of conditioned responding. *Psychological Research*. <https://doi.org/10.1007/s00426-020-01345-w>
- Zaman, Jonas, Madden, V. J., Iven, J., Wiech, K., Weltens, N., Ly, H. G., ... Van Diest, I. (2017). Biased Intensity Judgements of Visceral Sensations After Learning to Fear Visceral Stimuli: A Drift Diffusion Approach. *The Journal of Pain*, 18(10), 1197–1208. <https://doi.org/10.1016/j.jpain.2017.04.011>
- Zaman, Jonas, Wiech, K., Claes, N., Van Oudenhove, L., Van Diest, I., & Vlaeyen, J. W. S. (2018). The Influence of Pain-Related Expectations on Intensity Perception of Nonpainful Somatosensory Stimuli. *Psychosomatic Medicine*, 80(9), 836–844. <https://doi.org/10.1097/PSY.0000000000000586>
- Zaman, Jonas, Wiech, K., & Vlaeyen, J. W. S. (2019c). Perceptual Decision Parameters and Their Relation to Self-Reported Pain: A Drift Diffusion Account. *Journal of Pain*, 00(00), 1–24. <https://doi.org/10.1016/j.jpain.2019.06.009>
- Zenses, A. K., Lee, J. C., & Zaman, J. (2020). Differences in perceptual memory influence generalization patterns. *Psyarxiv*, 1–36. <https://doi.org/10.31234/osf.io/tc8f5>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.