Transfer of Conditioned Fear-potentiated Startle across Equivalence Classes. An Exploratory Study

Miguel Ángel López Medina, Miguel Rodríguez Valverde[*], Mónica Hernández López

Universidad de Jaén, España

Abstract

Research on fear conditioning is key to understanding the genesis and maintenance of anxiety disorders. A still scarce but growing evidence shows that fear-conditioned arousal reactions may transfer amongst physically dissimilar but symbolically related (e.g. equivalent) stimuli. The limited investigation published to date has relied on skin conductance responses as its main measure. Thus far, no published studies have analyzed this phenomenon using more emotionally sensitive psychophysiological measures, like fear-potentiated startle. Twenty-seven participants underwent a matching-to-sample procedure for the formation of two four-member equivalence classes (A1-B1-C1-D1 and A2-B2-C2-D2). Then, one element from each class was used in a differential aversive conditioning procedure (CS+: B1; CS-: B2) with electric shock as the UCS. Eye-blink startle (measured as EMG activity of the orbicularis oculi muscle after a burst of white noise), skin conductance responses, and shock-risk self-report ratings were collected. Results show no evidence of transfer of functions with any of the psychophysiological measures. A weak, inconclusive effect was observed for self-reported ratings.
Key words: transfer of functions, fear conditioning, generalization, equivalence relations, fear-potentiated startle.

Novelty and Significance

What is already known about the topic?

- Fear generalization is a key process for behavioral accounts of the anxiety disorders. There is limited but growing evidence that fear-conditioned arousal reactions do transfer across physically dissimilar arbitrary categories (equivalences classes).
- All research on transfer of fear has been conducted employing skin conductance responses as the measure of arousal. Fear-potentiated startle appears to be a much more adequate measure for the study of fear conditioning.
- Prior research has shown primary stimulus generalization of conditioned fear-potentiated startle in humans.

What this paper adds?

- This is the first study to explore transfer of conditioned fear with fear-potentiated startle as a measure.
- The study shows a strong aversive conditioning with self-report measures, but only a partial and inconclusive transfer effect.
- The study shows only weak evidence of aversive conditioning with fear-potentiated startle and an absence of transfer of function effects.

Conditioned fear, as well as conditioned avoidance, is a key element for behavioral accounts of anxiety and its disorders (Barlow, 2002). In fear conditioning, a neutral stimulus (CS) acquires the capacity to elicit fear (and related emotional reactions) after being repeatedly paired with an aversive unconditioned stimulus (UCS). However, for many instances of human fear in daily life it may be difficult to find an instance of aversive conditioning that accounts for the observed current fear (King, Gullone, & Ollendick, 1998; Menzies & Clarke, 1993a,b). Fear generalization to situations other

than the original aversive event appears as a key process for the explanation of these instances. Watson and Rayner's (1920) famous little Albert case constitutes the canonical example of fear generalization: after fear conditioning to a white rat, a toddler (little Albert) showed fear reactions to diverse furry things, none of which had been directly paired with aversive stimulation. This example illustrates the more traditional view of generalization as a form of transfer along a continuum of perceptual similarity amongst the directly conditioned stimulus and the generalization stimuli. This type of generalization of fear in humans is being researched with renovated interest in recent times (e.g., Dunsmoor & LaBar, 2013; Lissek, Biggs, Rabin, *et al*., 2008; Lissek, Powers, McClure, et al., 2014; van Meurs, Wiggert, Wicker, & Lissek, 2014).

Generalization can also take place on the basis of features other than physical similarity (Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015). Increasing evidence shows that fear reactions transfer according to conceptual/semantic categories. For instance, Dunsmoor, White, and LaBar (2011) found that fear generalized more easily between conceptually related stimuli (e.g. spider and web) than between conceptually unrelated stimuli (e.g. spider and hospital corridor). Research on this form of conceptual/ semantic transfer has the disadvantage of poor experimental control over the learning histories that might have resulted in the formation of the semantic categories to which the stimuli belong. For instance, words like web and spider must have been directly paired on uncountable occasions during the learner's lifetime, which renders conceptual/ semantic generalization findings amenable to explanation in terms of higher-order or sensory preconditioning (see Declerq & De Houwer, 2009).

Beyond semantic relatedness, a limited but growing empirical evidence shows that fear generalization also occurs between stimuli belonging to the same de novo created category (e.g. Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; Luciano, Valdivia Salas, Ruiz, *et al*., 2013; Rodríguez Valverde, Luciano, & Barnes-Holmes, 2009; Vervoort, Vervliet, Bennett, & Baeyens, 2014), that is, a set of physically dissimilar stimuli that are grouped arbitrarily upon some form of training. This type of work has been conducted in the field of behavior analysis under the labels stimulus equivalence and transfer of functions (Dougher & Markham, 1996; Hayes, Gifford, & Wilson, 1996; Sidman, 1994) (more recently this phenomenon has also been researched under the umbrella term symbolic generalization; see Dymond, Dunsmoor, Vervliet, Roche, & Hermans, 2015).

Stimulus equivalence emerges readily in language-able humans after training interrelated conditional discriminations among arbitrary (physically dissimilar) stimuli. Typically, conditional discriminations are trained and tested with matching-to sample (MTS) procedures: for instance, when choosing B1 in the presence of A1 is reinforced (A1-B1), and choosing C1 in the presence of A1 is also reinforced (A1-C1), untrained relations of symmetry (choosing A1 in the presence of either B1 or C1: B1-A1 and C1-A1) and transitivity (choosing C1 in the presence of B1, and vice versa: B1-C1 and C1-B1) will appear in the absence of feedback. When combined symmetry and transitivity relations emerge amongst a set of stimuli, a stimulus equivalence relation is said to have been established amongst these stimuli (Sidman & Tailby, 1982; Stewart & Roche, 2013). Research has repeatedly shown that the stimuli belonging to this type of

novel category (an equivalence class) are functionally substitutable, that is, when one of them acquires a function (i.e., is trained to control a new behavior), the remaining elements of the class will have the same function without explicit training and without having been experienced together (e.g., Augustson & Dougher, 1997; Dougher *et al*., 1994; Rodríguez Valverde *et al*., 2009).

In spite of its relevance for a behavioral account of anxiety disorders, very few studies have analyzed the transfer of aversively conditioned fear as measured by changes in autonomic arousal. Dougher *et al*., (1994) provided the first demonstration of transfer of conditioned fear among elements of the same equivalence class in a small study (*N*= 8). After training for the formation of two four-member equivalence classes (A1-B1-C1-D1 and A2-B2-C2-D2), participants underwent differential aversive conditioning with one element from each class (CS+: B1; CS-: B2). Then they were tested with other elements from each class (C1, D1, C2 and D2). Participants showed higher skin conductance responses (SCRs) with the elements belonging to the CS+ category (C1 and D1). Later research by Rodríguez Valverde *et al*. (2009), employing a larger sample and conditioning parameters and statistical methods more consistent with contemporary standards in human conditioning research, confirmed these findings and extended on them. This study included a contingency reversal phase that allowed for the within-subject replication of the transfer of fear effect. In brief, after the acquisition of differential conditioning and transfer tests, two elements from the original CS+ category were used as CS-, and two elements from the original CS- category were used as CS+. Then, the remaining elements from each category were tested, and they changed their function according to the latter contingency arrangement, proving that the observed transfer effects were not due to a methodological artifact. More recently, Vervoort *et al*. (2014) have investigated the transfer of conditioned fear using SCRs and online expectancy ratings as dependent variables. While this study showed partial evidence of the transfer of conditioned fear across members of the same category with shock-expectancies, the evidence regarding SCR was unclear (differences between the test stimuli from each class only approached significance; see Vervoort *et al*., 2014, p. 5).

All of the few published studies on the transfer of fear that have collected psychophysiological measures have used SCR as their main dependent variable. While SCR is one of the most popular psychophysiological measures, its validity as a measure of conditioned fear and anxiety has been questioned (Lissek *et al*., 2005). SCR reflects increases in general sympathetic arousal that can be associated to a number of different emotions other than fear and anxiety (Cacioppo, Berntson, Larsen, Poehlmann, & Ito, 2000). In addition, SCR is very sensitive to attentional processes (Filion, Dawson, Schell, & Hazlett, 1991; Frith & Allen, 1983; Gray & McNaughton, 2000), and may even be taken as to reflect attention rather than emotion (e.g., Carretié, Mercado, Tapia, & Hinojosa, 2001; Öhman, Flykt, & Esteves, 2001). Recent research is increasingly relying on fear-potentiated startle (FPS: the enhancement of the startle reflex upon a state of fear, Grillon, 2008) as a valid psychophysiological measure of fear (Lissek *et al*., 2008). FPS is more specific to the affective valence of stimuli, as proved by research on fear potentiation and attenuation (e.g., Bradley, Cuthbert, & Lang, 1999). Also, attentional effects on FPS seem to be smaller than emotional effects (e.g. Bocker,

Baas, Kenemans, & Verbaten, 2004). Finally, it appears to be difficult to consciously influence FPS, as it is a brainstem reflex mediated by a small number of synapses (Lipp, 2006). Given these features, FPS is being increasingly used as a psychophysiological proxy of pathological anxiety (Grillon, 2008).

Although, as far as we know, there is no published research exploring the transfer of fear with FPS, this measure has been employed to investigate primary stimulus generalization. Lissek et al. (2008) conducted an aversive differential conditioning procedure with 10 circles of increasing size (and identical in all other respects) as stimuli. The extremes of this stimulus continuum (the smallest and the largest circle) served either as CS+ or CS- with mild electric shock as the UCS. The eight circles of intermediate size served as generalization stimuli in a continuum of similarity from the CS+ to the CS-. Lissek and colleagues collected both FPS and self-reported shock-expectancy ratings. Both measures showed a gradual weakening of the conditioned-fear response as the tested stimulus became less similar to the CS+ (i.e., a generalization gradient).

The present study investigates the transfer of conditioned-fear functions using FPS as a measure. The experimental procedure employed here is a combination of that in Rodríguez Valverde *et al*. (2009, Experiment 1) with an adaptation of the aversive conditioning procedure employed in Lissek *et al*. (2008). In brief, participants were trained in the formation of two four-member equivalence classes (A1-B1-C1-D1 and A2-B2-C2-D2). Subsequently, element B1 served as CS+ in a differential aversive conditioning procedure, while element B2 served as CS-. Mild electric shock was used as the UCS. Subsequently, elements C1, D1, C2 and D2 were tested for transfer. In addition to FPS, SCR and online self-reported shock expectancy were measured.

Method

Participants

Participants were 27 undergraduates (from Universidad de Jaén, España; 20 female) with a mean age of 20.04 (age range 18-35). Participants were recruited through in-class announcements and received course credit as compensation for participation. At the beginning of the experiment participants read and signed a statement of informed consent (approved, like all the experimental procedures by the Ethics Research Board). Upon termination of the experiment, all participants were debriefed.

Setting, Apparatus, and Stimuli

The experiment was conducted in a laboratory consisting of two adjacent rooms (an experimental room and an observation room) separate by a one-way mirror for participant observation. The experimental room was equipped with a table, an armchair, a Windows compatible desktop computer (15-inch color screen, with standard mouse and keyboard), and a computerized physiological recording system (BIOPAC Instruments MP150 with GSR100C and EMG100C modules). The computer was programmed (with VisualBasic® 6.0) to present visual stimuli and to coordinate the presentation of electric shocks, as well as to record participants' responses on some tasks. The observation room

contained a desktop computer for the collection, storage and analysis of physiological response data.

Startle-blink EMG was recorded with two Ag/AgCl reusable electrodes (4-mm diameter and 2 mm-deep gel cavity) filled with standard electrolyte (SignaGel) placed under the left eye (muscle orbicularis oculi). Startle was elicited by a white-noise burst (102 dB, 50-ms duration) presented binaurally through headphones. Skin conductance (SC) was measured through non-polarizable disposable Ag/AgCl electrodes (11-mm contact surface) attached to the thenar and hypothenar eminences of the participant's left hand. The system delivered a constant voltage (0.5 V) current that allowed for the continuous measurement of tonic skin conductance level (DC) (i.e. constant voltage technique of exosomatic recording; see Dawson, Schell, & Filion, 1990). Analog-digital conversion was performed at 2000 Hz. An isolated square-wave stimulator was used for the administration of constant voltage electric shocks (50 ms duration) through two disposable adhesive round electrodes (15 mm diameter and 40 mm apart) attached to the inner surface of the participant's left arm. The computer that presented the experimental tasks was connected to the physiological recorder and the stimulator through a custom-made parallel-port cable that delivered the electric pulses that served as stimulus markers for the physiological records and that triggered the presentation of shocks.

The visual stimuli were 12 black abstract shapes (the same as in Dougher *et al*., 1994, p. 333, and Rodríguez Valverde *et al*., 2009) framed in a square white background, presented on a general black background. The size of the stimuli was 8×8 cm^2. For each participant, the shapes were randomly distributed into three four-member groups by the computer. The stimuli in each group were designated with alphanumerical labels (e.g. A1, B2, C3, etc.) for procedural purposes. The participants were unaware of these labels.

Online shock expectation rating was collected through the presentation of the question "Level of risk?" on the computer screen above the visual stimulus two seconds after trial onset. Participants completed this rating according to a three-point Likert scale (similarly to Lissek *et al*., 2008, p. 680) (1= no risk, 2= moderate risk, and 3= high risk). Participants had to press buttons 1, 2, or 3 on the computer keyboard as quickly as possible upon the presentation of the risk-level question. As in Lissek *et al*. (2008), expectancy ratings were elicited on half of the trials, and startle responses were probed on half of the trials. This separation of expectancy ratings and startle responses was implemented in order to prevent that ratings would interfere with physiological responses (see Lissek *et al*., 2008, pp. 680-681).

Procedure

The experiment comprised four different phases. Phase 1 consisted of the formation (training and evaluation) of two four-member equivalence classes (A1-B1-C1-D1 and A2-B2-C2-D2). Phase 2 was a pre-acquisition procedure, where B1, B2, the startle-inducing white noise burst, and the shock expectancy rating were presented for habituation. Phase 3 consisted of an aversive differential conditioning procedure, where electric shocks of moderate intensity were used as unconditional stimulation (UCS). B1 served as a CS+ (was paired with shock) and B2 served as a CS- (was never paired

with shock). Phase 4 consisted of a transfer test of respondent elicitation to the other elements of each class (C1, D1, C2, and D2). Phases 2 to 4 were conducted as part of the same experimental task, without stops or breaks amongst them. During this task, visual stimuli were individually presented in the center of screen with a fixed 6-s duration, with an inter-trial interval (ITI) of random duration between 14 and 18 s. In half of the trials, startle was probed by presenting the white-noise burst 4 s after visual stimulus onset, and in the other half, the question "Level of risk?" was presented 2 s after stimulus onset (see below).

All phases were conducted in one experimental session that lasted between 90 and 120 min. Participants were run individually.

Upon reporting to the laboratory, participants were interviewed and were explained the general procedures that would follow. The experimenter placed special emphasis on the fact that participation was absolutely voluntary, and that participants were free to abandon the experiment at any time. He also told them that participation entailed receiving the administration of several shocks of moderate intensity, the level of which would be selected by the participants themselves, so that the sensation produced by the shock was definitely unpleasant, but not painful. After that, participants read and signed a statement of informed consent that, in addition to the points already mentioned by the experimenter, explicitly highlighted that participants suffering from any sort of cardiovascular disease, epilepsy, or any serious illness should not take part in the study.

Phase 1: Formation of equivalence classes. During this phase, participants underwent training in six different conditional discriminations (A1B1, A1C1, A1D1, A2B2, A2C2, and A2D2) through a simultaneous one-to-many matching-to-sample procedure (Sidman & Tailby, 1982) with three comparisons. After the establishment of these six relations, they underwent the assessment of derived symmetry and transitivity relations, in order to determine whether two equivalence classes had been established (Class 1: A1-B1-C1-D1; Class 2: A2-B2-C2-D2). One set of abstract shapes (A3, B3, C3, and D3) served as incorrect comparisons in conditional discrimination training and testing. No specific relations were trained nor tested for this set of stimuli. This was done because the interest was in teaching two equivalence classes, not three, but it was also important to have three comparisons in each trial, in order to control for responding by exclusion (see Carrigan & Sidman, 1992). This procedure is similar to that in previous studies on transfer of aversive respondents (Dougher *et al*., 1994; Rodríguez Valverde *et al*., 2009).

Both during training and test trials, a sample (e.g. A1) appeared in the center of the upper third of the screen. One second later, three comparisons (e.g. B1, B2 and B3) appeared in line in the horizontal lower third of the screen. On each trial, each comparison's position (left, center, or right) was randomly assigned. Participants responded by selecting one of the three comparisons with a mouse-click, after which all the stimuli were removed from the screen. During training trials, correct responses were followed by the word "BIEN" (i.e. good) in capital letters and white color, centered on the screen. Incorrect responses were followed by the word "MAL" (i.e. wrong) in the same format. Feedback remained on the screen for 1.25 s, after which the screen went blank for 1 s. During test trials, no feedback was presented. The following instructions were presented on the screen at the start of this phase:

"In this part of the experiment you will see four shapes on the screen in each trial, one in the middle of the top of the screen, and the other three at the bottom: one on the right, one on the left, and one in the middle. The task consists of selecting the correct shape out of those three in the lower part of the screen, by clicking on it with the mouse. During the first part of the task, the computer will give you feedback on each response, indicating whether it is correct (BIEN) or wrong (MAL). Later in the task, the computer will not give feedback about your responses. However, there will be correct and incorrect responses, and you should do your best to have as many correct responses as possible.

At first the task will be easy, and you may even notice that it is not necessary to pay a lot of attention. However, task difficulty will increase gradually, and in order to select the correct shapes during the last part of the task, you must have previously performed the task correctly during the initial parts. So, it is important that you pay a lot of attention from the beginning of the task. Everything you learn during this part of the experiment will be important for later phases. If you have any doubts, please ask the experimenter. Click here to start."

The six directly trained relations were presented in six trial blocks (one trial per relation). Each trial consisted of the presentation of the sample and its corresponding group of comparisons. Within each block, the presentation order of trials was randomized. Blocks were presented continuously until the participant achieved a mastery criterion of 34 correct responses out of 6 consecutive complete blocks (36 trials). After that, participants underwent a test of symmetry relations, preceded by the following instructions:

"Now you will keep performing the task. This time the computer will not give you any feedback about your responses, but there are still correct and incorrect responses. You have to get as many correct responses as possible. Click here to continue."

The six tested symmetry relations (B1A1, B2A2, C1A1, C2A2, D1A1, and D2A2) were continuously presented in six-trial blocks (one trial per relation in random order within each block) until participants reached the mastery criterion (34 correct responses out of 6 consecutive blocks), up to a maximum of five six-block sets (i.e. 240 trials). In case the criterion was not achieved, participants were scheduled for retraining the next day.

Upon criterion achievement, participants underwent a final combined test of symmetry and equivalence relations, preceded by the same instructions as for the symmetry test. This test comprised the 6 symmetry relations and 12 equivalence relations (i.e. B1A1, B2A2, C1A1, C2A2, D1A1, D2A2, B1C1, B2C2, B1D1, B2D2, C1D1, C2D2, C1B1, C2B2, D1C1, D2C2, D1B1, and D2B2). Both types of relation were presented continuously in 18-trial blocks (one trial per relation in random order) until achievement of the mastery criterion of 85 correct responses out of six consecutive blocks (i.e. 90 trials), with a maximum of two errors for each particular relation, up to a maximum of five five-block sets (i.e. 450 trials). In case the criterion was not achieved, participants were scheduled for retraining the next day. After that, participants passed to the next phase of the experiment.

Phase 2: Pre-acquisition. This phase began with participant preparation for psychophysiological recording. All electrodes (stimulation and recording) were attached as indicated above (see Setting, Apparatus and Stimuli). Then, participants underwent a shock workup procedure wherein they selected the intensity of the shock that would be used as UCS during the conditioning procedure (to a point perceived as definitely unpleasant, though not painful). After the shock workup procedure, the following instructions were presented on the screen:

"During this phase of the experiment you will be presented with some shapes on the screen, one at a time. It is important that you pay attention to the screen and to the shapes appearing on it. On some occasions, you will receive shocks of the same magnitude you have just selected. Initially the computer screen will remain blank for a few minutes, in order to have your physiological activity at a steady level. After this interval, shapes will begin to appear on the screen. All you have to do is pay attention to the screen, and remain seated and quiet. It is important that you try not to move, cough, sneeze, laugh, etc., as all of these actions may interfere with the physiological recordings. On some occasions, the question "Level of risk?" will be presented on the screen, asking for the risk of receiving a shock in that precise moment. You have to respond as quickly as possible by pressing the correct key on the keyboard (1=no risk; 2=moderate risk; 3=high risk) with your right index finger. Other than that, it is important that you remain seated and quiet.

If you have any doubts, please ask the experimenter. Otherwise, press the space bar to continue."

Once the participant confirmed that s/he had understood the instructions, the experimenter left the room and the procedure started. There was a 10-min baseline period during which the screen remained blank. Then, the presentation of stimuli began. Participants were first presented with the white-burst noise on four consecutive occasions for habituation. Then, they were presented with the "Level of risk?" question four consecutive times (ITI= 12 s). Finally, they were presented with stimuli B1 and B2. Each stimulus was presented twice (once for shock-expectancy rating and once for startle probing) in random order. During this phase, neither stimulus was paired with shock.

Phase 3: Conditioning acquisition. This phase consisted of an aversive differential delay conditioning procedure, where B1 served as CS+ and B2 served as CS-. Each CS was presented 10 times on three-trial blocks (one CS+, one CS-, and one blank-screen ITI-trial per block). On odd blocks, startle was probed with the presentation of the white-noise burst 4 s after CS (or ITI-trial) onset. On even blocks, shock-expectancy rating was assessed with the presentation of the question 2 s after CS (or ITI-trial) onset. The offset of B1 was simultaneous to the presentation of electric shock (UCS onset) on 80% of B1 trials. Shock was never presented on B2 or ITI trials.

Phase 4: Transfer of function tests. Stimulus presentation was conducted in a very similar manner to Phase 3, with the difference that, in addition to B1 (CS+), B2 (CS-) and the ITI-trials, the following stimuli were presented: C1, D1, C2, and D2 (i.e., all the elements of each class were presented with the exception of A1 and A2, as the responses to these two elements could be interpreted in terms of higher-order conditioning, rather than in terms of transfer of functions: see Dymond & Rehfeldt, 2000). Each stimulus was presented 8 times (i.e., eight 7-trial blocks, one trial per stimulus in random order). As in the previous phase, startle was probed on odd blocks and shock-expectancy rating was measured on even blocks. On 50% of B1 trials, the offset of B1 was simultaneous to the presentation of electric shock (UCS onset). This was done in order to prevent extinction of directly conditioned fear to B1. No other stimulus was paired with shock.

Response Quantification and data analysis

Startle EMG was rectified and smoothed based on a 20-ms window average. In each startle-probe trial, peak amplitude was determined for responses occurring within

a 150-ms window after startle-noise (with a 20-ms minimum latency onset window). The average EMG level of a 50-ms baseline immediately preceding startle-noise onset was subtracted from this peak amplitude. EMG amplitude scores were standardized into T-scores (within subject) in order to normalize data and reduce between-subject variability.

Anticipatory SCRs were quantified as follows: for each visual stimulus presented during startle probe trials in each phase, the largest increase in skin conductance level (SCL, measured in *μSiemens* [*μS*]) was calculated during the 4-s interval between stimulus onset and startle-noise onset. This variation was measured from the point of response onset (minimum latency window 0.5 s) to the highest SCL value within the 4-s interval. SCL decreases or SCL increases starting prior to the 0.5-s minimum latency window were quantified as zero.

Each measure (startle, SCR, and shock-expectancy) was analyzed individually with a repeated measures (RM) MANOVA in each phase, with stimulus type as the RM factor (CS+, CS- in pre-acquisition and acquisition; CS+, CS-, Class 1, Class 2) in the transfer test). MANOVAs were computed using Wilk's Lambda. Where necessary subsequent paired samples t tests were performed for post-hoc planned comparisons.

RESULTS

Table 1 presents descriptive statistics for all three measures in phases 2, 3 and 4. As expected, for Phase 2 (pre-acquisition, prior to aversive conditioning), there was no main effect for CS-type for any of the three measures (all $ps >.05$), with shock-expectancy approaching significance, $F_{(1, 26)}= 3.566$, $p= .069$, $\eta^2p= .121$. During the conditioning acquisition phase, the only measure where differential conditioning was observed was shock expectancy, with significantly higher expectation ratings for B1 (B1/CS+: $M= 2.390$, $SD= 0.570$; B2/CS-: $M= 1.61$; $SD= .59$), $F_{(1, 26)}= 26.144$, $p <.001$,
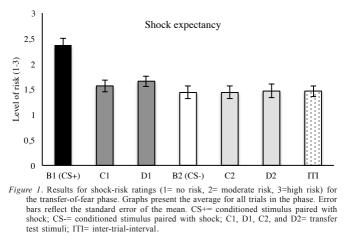
Table 1. Descriptive statistics for startle EMG (T scores), shock-expectancy rating (1-3), and skin conductance responses across phases 2, 3, and 4.

Stimulus	Startle		Shock- expectancy		SCR	
	Phase 2: Pre-acquisition					
	M	SD	M	SD	M	SD
B1	54.16	7.45	1.48	0.47	0.28	0.34
B2	57.61	8.75	1.70	0.76	0.30	0.36
	Phase 3: Conditioning acquisition					
Stimulus	Startle		Shock expectancy		SCR	
	M	SD	M	SD	M	SD
B1	54.23	5.23	2.39	0.57	0.17	0.28
B2	52.23	4.40	1.61	0.59	0.13	0.17
ITI	48.07	3.46	1.73	0.61	--	--
	Phase 4: Transfer test					
Stimulus	Startle		Shock expectancy		SCR	
	M	SD	M	SD	M	SD
B1	51.36	5.76	2.37	0.69	0.18	0.28
B2	47.23	3.99	1.44	0.64	0.06	0.14
C1	46.50	5.66	1.56	0.61	0.09	0.14
C2	49.79	4.49	1.44	0.67	0.07	0.12
D1	49.95	4.23	1.66	0.54	0.09	0.15
D2	47.97	4.78	1.46	0.69	0.11	0.23
ITI	45.65	5.33	1.46	0.57	--	--

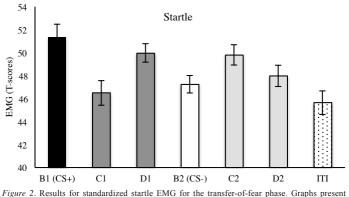*Notes: M=* mean; *SD=* Standard Deviation.

$\eta^2 p$= .501 (EMG startle: p= .116; SCR: p= .880).

During the transfer test phase there was a clear stimulus-type effect for shock-expectancy, $F_{(3, 24)}$= 8.674, p= .001, $\eta^2 p$= .520 (see Figure 1).

Figure 1. Results for shock-risk ratings (1= no risk, 2= moderate risk, 3=high risk) for the transfer-of-fear phase. Graphs present the average for all trials in the phase. Error bars reflect the standard error of the mean. CS+= conditioned stimulus paired with shock; CS-= conditioned stimulus paired with shock; C1, D1, C2, and D2= transfer test stimuli; ITI= inter-trial-interval.

Planned comparisons among tested stimuli showed that shock-expectancy rating was significantly higher for B1 (*M*= 2.37, *SD*= .69) than for B2 (*M*= 1.44, *SD*= .64), $t_{(26)}$= 5.272, p <.001, *d*= 1.02, which is clearly indicative of differential conditioning with this measure. A planned comparison between averaged C1 and D1 (*M*= 1.61; *SD*= .55), on the one hand, and C2 and D2 (*M*= 1.45; *SD*= .67), on the other, was significant too [$t_{(26)}$= 2.384, *p*= .025, *d*= .48], which is apparently compatible with a transfer of function effect. However, this transfer effect would be incomplete, given that the difference between B1 and B2 was significantly larger than the difference between C1D1 and C2D2, $t_{(26)}$= 4.550, *p* <.001, *d*= 1.03.
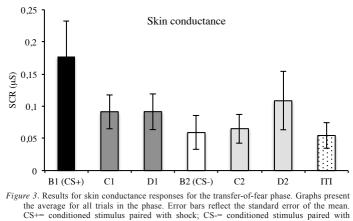
During the transfer test phase there was no significant stimulus-type effect either for startle EMG (see Figure 2) or SCR (see Figure 3), although it approached significance
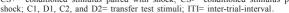

Figure 2. Results for standardized startle EMG for the transfer-of-fear phase. Graphs present the average for all trials in the phase. Error bars reflect the standard error of the mean. CS+= conditioned stimulus paired with shock; CS-= conditioned stimulus paired with shock; C1, D1, C2, and D2= transfer test stimuli; ITI= inter-trial-interval.

Figure 3. Results for skin conductance responses for the transfer-of-fear phase. Graphs present the average for all trials in the phase. Error bars reflect the standard error of the mean. CS+= conditioned stimulus paired with shock; CS-= conditioned stimulus paired with shock; C1, D1, C2, and D2= transfer test stimuli; ITI= inter-trial-interval.

in both cases [startle EMG: $F_{(3,24)}= 2.905$, $p= .055$, $\eta^2p= .266$; SCR: $F_{(3,24)}= 2.591$, $p= .076$, $\eta^2p= .245$]. For startle EMG, planned comparisons between B1 ($M= 51.36$, $SD= 5.76$) and B2 ($M= 47.23$, $SD= 3.99$) showed a significant difference indicative of differential conditioning, $t_{(26)}= 2.988$, $p= .006$, d= .58. However, there was no difference between averaged startle to C1D1 and to C2D2 [$t_{(26)}= .741$, $p= .465$], which is clearly indicative of an absence of a transfer of function effect. The same pattern was observed for planned comparisons with SCR as a measure, with a significant difference [$t_{(26)}= 2.748$, $p= .020$, $d= .59$] between B1 ($M= .24$, $SD= .40$) and B2 ($M= .06$, $SD= .13$), and no difference between C1D1 and C2D2 [$t_{(26)}= -.412$; $p= .684$].

Discussion

The present study is, to our knowledge, the first attempt to explore the transfer of conditioned fear across equivalence classes using fear-potentiated startle (FPS) as a measure of fear-conditioned arousal. The results showed no evidence of such transfer effect, neither with FPS nor with SCR. In principle, it seems that a possible explanation for this absence of transfer of fear with arousal measures might be weak aversive conditioning. During the acquisition phase, no conditioning was observed with either physiological measure. Given that during the transfer phase (wherein B1 was paired with shock on 50% of the trials) there was a clear differential conditioning effect (with B1 eliciting stronger FPS and SCR than B2), it appears that the amount of conditioning acquisition trials was just too small to guarantee the conditioning effect during acquisition. Perhaps a longer acquisition phase might have produced stronger aversive conditioning and a higher likelihood of transfer of functions during test trials.

The absence of transfer effects might also be attributed to a problem in the maintenance of equivalence relations, although in principle this seems unlikely considering how easily all participants reached strict training and test criteria for class formation.

Due to procedural concerns (the experimental sessions were lengthy) this study did not comprise a re-test of the equivalence classes trained and assessed in Phase 1. Other studies on transfer of functions incorporate such re-test (e.g. Rodríguez Valverde et al., 2009) as a means to ascertain that the equivalence relations are maintained throughout the experimental procedures.

In any case, the fact that derived relations are maintained in a re-test (i.e., the same context where they were trained and tested, the MTS procedure) does not necessarily entail that they will be relevant during transfer probe trials. Given that transfer test trials (with C1, D1, C2, and D2) were conducted in extinction, the arranged contingencies during this test might have conflicted with previous contingencies according to which members of Class 1 (B1) had reliably predicted the presentation of shock. After repeated pairing of B1 with shock and B2 with absence of shock, the fact that C1 and D1 were presented in extinction conflicted with the functional equivalence that would be expected from the previously established equivalence relations. This may have had the effect of establishing that the conditioning-transfer task was a context where the previously established equivalence relations were not relevant for the prediction of shock presentation. Previous research from our own group with SCR as the main measure presented a similar limitation (see Rodríguez Valverde et al., 2009, Experiment 1). This problem was addressed in the present study with the use of a partial CS–UCS contingency during acquisition in order to attenuate the effect of extinction (like in Lissek et al., 2008; see also Roche & Barnes, 1997; Vansteenwegen, Crombez, Baeyens, Hermans, & Eelen, 2000). It is obvious that this did not work to guarantee that FPS and SCR to elements of Class 1 were consistently larger than to elements of Class 2 during the transfer test. Future research should consider the use of alternative methodological solutions, like presenting transfer test stimuli with the same aversive conditioning contingencies as their class counterparts during conditioning acquisition, and then take the first presentation of each transfer stimulus as the only valid test probe. This solution proved useful for the obtention of a solid transfer of fear effect in prior research with SCR as the main measure (see Rodriguez Valverde et al., 2009, Experiment 2). In any case, an analysis of transfer effects based on the first presentation of test stimuli (not reported above) did not yield any different outcomes in the present study.

A slightly different pattern of results emerged from the analysis of shock-expectancy data. Results with this measure were indicative of a clear aversive conditioning effect (both during conditioning acquisition and during the transfer test). Also, although incomplete, there is some indication of a transfer effect, with consistently higher shock-expectancy ratings to C1D1 than to C2D2. However, as mentioned above, this difference was much smaller than the observed difference between B1 and B2. This pattern of results, with stronger fear conditioning for expectancy than for arousal measures, and a partial transfer effect for the former, is not unlike previous findings in the area (see Vervoort et al., 2014). The lack of correlation among different measures of fear (arousal, subjective experience, and action-tendency/avoidance) has been pointed out as indicative of the multidimensionality of fear expression, an issue that is insufficiently addressed in fear conditioning research (see Beckers, Krypotos, Boddez, Effting, & Kindt, 2013). Besides, it could be questioned whether the shock-expectancy ratings in this study were

a measure of fear at all, or merely and index of awareness of the aversive conditioning contingency.

All in all, this experiment failed to find transfer of fear-conditioned arousal with FPS and SCR as measures. The aversive conditioning procedure here employed was an adaptation of a procedure that served to find perceptual fear generalization with FPS (Lissek *et al.*, 2008), but it did not prove useful for the observation of transfer of conditioned fear after the successful formation of two different categories (i.e. equivalence classes). This may be indicative that the latter form of fear generalization (i.e. conceptual or symbolic generalization) is more complex in nature, and further research with methodological modifications will be necessary in order to see if it is possible at all to observe transfer of function effects with FPS. Related to this, there is a contrast between the scarcity of articles reporting transfer-of-fear effects with arousal measures, and the relative abundance of reports presenting evidence for the transfer of avoidance-evoking functions (see Augustson & Dougher, 1997; Dymond, Roche, Forsyth, Whelan, & Rhoden, 2007, 2008; Dymond, Schlund, Roche, De Houwer, & Freegard, 2012; Dymond, Schlund, Roche, & Whelan, 2014; Dymond *et al.*, 2011; Gannon, Roche, Kanter, Forsyth, & Linehan, 2011; Luciano *et al.*, 2013; Roche, Kanter, Brown, Dymond, & Fogarty, 2008). Although this may just show a higher difficulty in experimentally controlling variability inherent to arousal measures, it may also be indicative of the relative importance of the avoidance component compared to the arousal one in human fear learning. Previous research has shown clear differences between arousal and action-tendency (i.e. avoidance) expressions of conditioned fear and transfer, with stimuli evoking strong avoidance responses (either directly conditioned or derived) in the absence of noticeable arousal (see Luciano *et al.*, 2013).

As mentioned above, FPS is viewed as an especially adequate psychophysiological measure of fear (e.g., Lipp, 2006) and a potentially powerful translational tool for the experimental study of a clinically relevant features of pathological anxiety (Grillon, 2008). The study of transfer of functions (i.e., the generalization of fear across an arbitrary conceptual, rather than a perceptual dimension) with this measure seems worthy of further exploration.

REFERENCES

Augustson EM & Dougher MJ (1997). The transfer of avoidance evoking functions through stimulus equivalence classes. *Journal of Behavior Therapy and Experimental Psychiatry, 28*, 181-191.
Barlow DH (2002). *Anxiety and its disorders: The nature and treatment of anxiety and panic, 2nd ed*. New York, NY: Guilford.
Beckers T, Krypotos AM, Boddez Y, Effting M, & Kindt M (2013). What's wrong with fear conditioning? *Biological Psychology, 92*, 90-96. Doi: 10.1016/j.biopsycho.2011.12.015.
Bocker KBE, Baas JMP, Leon Kenemans J, & Verbaten MN (2004). Differences in startle modulation during instructed threat and selective attention. *Biological Psychology, 67*, 343-358.
Bradley MM, Cuthbert BN, & Lang PJ (1999). Affect and the startle reflex. In ME Dawson & AM Schell (Eds.), *Startle modification: Implications for neuroscience, cognitive science, and clinical science* (pp. 157-183). Cambridge: Cambridge University Press.

Cacioppo JT, Berntson GG, Larsen JT, Poehlmann KM, & Ito TA (2000). The psychophysiology of emotion. In M Lewis & JM Haviland-Jones (Eds.), *Handbook of emotions*, *2nd ed*. (pp. 173-191). New York: Guilford Press.

Carretié L, Mercado F, Tapia M, & Hinojosa J (2001). Emotion, attention, and the 'negativity bias', studied through event-related potentials. *International Journal of Psychophysiology, 41*, 75-85.

Carrigan PF & Sidman M (1992). Conditional discrimination and equivalence relations: A theoretical analysis of control by negative stimuli. *Journal of the Experimental Analysis of Behavior, 58*, 183-204.

Dawson ME, Schell AM, & Filion DL (1990). The electrodermal system. In JT Cacioppo & LG Tassinary (Eds.), *Principles of Psychophysiology* (pp. 295-324). New York: Cambridge University Press.

Declerq M & De Houwer J (2009). Transfer of avoidance responding to a sensory preconditioned cue: Evidence for the role of S-S and R-S knowledge in avoidance learning. *Learning and Motivation, 40*, 197-208. Doi: 10.1016/j.lmot.2008.11.003

Dougher MJ & Markham MR (1996). Stimulus classes and the untrained acquisition of stimulus functions. In TR Zentall & PM Smeets (Eds.) *Stimulus Class Formation in Humans and Animals* (pp. 71-90). Amsterdam: North Holland.

Dougher MJ, Augustson E, Markham MR, Greenway DE, & Wulfert E (1994). The transfer of respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 62*, 331-351.

Dunsmoor J & LaBar K (2013). Effects of discrimination training on fear generalization gradients and perceptual classification in humans. *Behavioral Neuroscience, 127*, 350-356. Doi: 10.1037/a0031933

Dunsmoor J, White A, & LaBar K (2011). Conceptual similarity promotes generalization of higher order fear learning. *Learning & Memory, 18*, 156-160. Doi: 10.1101/lm.2016411.

Dymond S & Rehfeldt RA (2000). Understanding complex behavior: The transformation of stimulus functions. *The Behavior Analyst, 23*, 239-254.

Dymond S, Dunsmoor JE, Vervliet B, Roche B, & Hermans D (2015). Fear Generalization in Humans: Systematic Review and Implications for Anxiety Disorder Research. *Behavior Therapy, 46*, 561-582.

Dymond S, Roche B, Forsyth JP, Whelan R, & Rhoden J (2007). Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *Journal of the Experimental Analysis of Behavior, 88*, 249-262.

Dymond S, Roche B, Forsyth JP, Whelan R, & Rhoden J (2008). Derived avoidance learning: Transformation of avoidance response functions in accordance with the relational frames of same and opposite. *The Psychological Record, 58*, 271-288.

Dymond S, Schlund MW, Roche B, De Houwer J, & Freegard GP (2012). Safe from harm: Learned, instructed, and symbolic generalization pathways of human threat-avoidance. *PLoS ONE 7(10)*: e47539. Doi: 10.1371/journal.pone.0047539.

Dymond S, Schlund MW, Roche B, & Whelan R (2014). The spread of fear: Symbolic generalization mediates graded threat-avoidance in specific phobia. *The Quarterly Journal of Experimental Psychology, 67*, 247-259. Doi: 10.1080/17470218.2013.800124

Dymond S, Schlund MW, Roche B, Whelan R, Richards J, & Davies C (2011). Inferred threat and safety: Symbolic generalization of human avoidance learning. *Behaviour Research and Therapy, 49*, 614-621. Doi: 10.1016/j.brat.2011.06.007

Gannon S, Roche B, Kanter J, Forsyth JP, & Linehan C (2011). A derived relations analysis of approach-avoidance conflict: Implications for the behavioural analysis of human anxiety. *The Psychological Record, 61*, 227-252.

Filion DL, Dawson ME, Schell AM, & Hazlett EA (1991). The relationship between skin conductance orienting and the allocation of processing resources. *Psychophysiology, 28*, 410-424.

Frith CD & Allen HA (1983). The skin conductance orienting response as an index of attention. *Biological Psychology, 17*, 27-39.

Gray JA & McNaughton N (2000). *The neuropsychology of anxiety*. New York: Oxford University Press.

Grillon C (2008). Models and mechanims of anxiety: Evidence from startle studies. *Psycopharmacology,*

199, 421-437.

Hayes SC, Gifford EV, & Wilson KG (1996). Stimulus classes and stimulus relations: Arbitrarily applicable relational responding as an operant. In TR Zentall & PM Smeets (Eds.) *Stimulus class formation in humans and animals* (pp. 279-300). Amsterdam: Elsevier.

King N, Gullone E, & Ollendick TH (1998). Etiology of childhood phobias: current status of Rachman's three pathways theory. *Behaviour Research and Therapy, 36*, 297-309.

Lipp OV (2006). Human fear learning: contemporary procedures and measurement. In MG Craske, D Hermans, & D Vansteenwegen (Eds). *Fear and Learning: from basic processes to clinical implications* (pp. 37-51). Washington DC: American Psychological Association.

Lissek S, Biggs AL, Rabin SJ, Cornwell BR, Álvarez RP, Pine DS, & Grillon C (2008). Generalization of conditioned fear-potentiated startle in humans: Experimental validation and clinical relevance. *Behaviour Research and Therapy, 46*, 678-687.

Lissek S, Bradford DE, Alvarez RP, Burton P, Espensen-Sturges T, Reynolds RC, & Grillon C (2014). Neural substrates of classically conditioned fear generalization in humans: A parametric fMRI study. *Social Cognitive and Affective Neuroscience, 9,* 1134-1142.

Lissek S, Powers AS, McClure EB, Phelps EA, Woldehawariat G, Grillon C, & Paine DS (2005). Classical fear conditioning in the anxiety disorders: A meta-analysis. *Behaviour Research and Therapy, 43,* 1391-1424.

Luciano C, Valdivia Salas S, Ruiz FJ, Rodríguez-Valverde M, Barnes-Holmes D, Dougher MJ, Cabello F, Vanessa Sánchez V, Barnes-Holmes Y, & Gutiérrez O (2013). Extinction of aversive eliciting functions as an analog of exposure to eliciting fear: Does it undermine avoidance responding? *Journal of Contextual Behavioral Science, 2,* 120-134. Doi: 10.1016/j.jcbs.2013.05.001.

Menzies RG & Clarke JC (1993a). The etiology of childhood water phobia. *Behaviour Research and Therapy, 31*, 499-501.

Menzies RG & Clarke JC (1993b). The etiology of fear of heights and its relationship to severity and individual response patterns. *Behaviour Research and Therapy, 31*, 355-365.

Öhman A, Flykt A, & Esteves F (2001). Emotion drives attention: Detecting the snake in the grass. *Journal of Experimental Psychology: General, 130*, 466-478.

Roche B, Kanter JW, Brown KR, Dymond S & Fogarty CC (2008). A comparison of 'direct' versus 'derived' extinction of avoidance. *The Psychological Record, 58*, 443-464.

Rodríguez Valverde M, Luciano C, & Barnes-Holmes D (2009). Transfer of aversive respondent elicitation in accordance with equivalence relations. *Journal of the Experimental Analysis of Behavior, 92*, 85-111.

Sidman M (1994). *Equivalence relations and behavior: A research story*. Boston: Authors Cooperative.

Sidman M & Tailby W (1982). Conditional discrimination versus matching-to-sample: An expansion of the testing paradigm. *Journal of the Experimental Analysis of Behavior, 37*, 5-22.

Stewart I & Roche B (2013). Relational Frame Theory: An overview. In S Dymond & B Roche (Eds.) *Advances in Relational Frame Theory: Research and applications* (pp. 51-71). Oakland, CA: New Harbinger/Context Press.

van Meurs B, Wiggert N, Wicker I & Lissek S (2014). Maladaptive behavioral consequences of conditioned fear generalization: A pronounced, yet sparsely studied, feature of anxiety pathology. *Behaviour Research and Therapy, 57*, 29-37. Doi: 10.1016/j.brat.2014.03.009

Vansteenwegen D, Crombez G, Baeyens F, Hermans D, & Eelen P (2000). Pre-extinction of sensory preconditioned electrodermal activity. *Quarterly Journal of Experimental Psychology, Section B: Comparative and Physiological Psychology, 53B*, 359-371.

Vervoort E, Vervliet B, Bennett M & Baeyens F (2014) Generalization of Human Fear Acquisition and Extinction within a Novel Arbitrary Stimulus Category. *PLoS ONE 9 (5)*: e96569. Doi: 10.1371/journal.pone.0096569

Watson JB & Rayner R (1920). Conditioned emotional reactions. *Journal of Experimental Psychology, 3*, 1-22.