

'Prepared' fear or socio-cultural learning? Fear conditioned to guns, snakes, and spiders  
is eliminated by instructed extinction in a within-participants differential fear  
conditioning paradigm

Camilla C. Luck, Rachel R. Patterson, & Ottmar V. Lipp

School of Psychology, Curtin University, Australia

Author Notes

Address for correspondence:

Camilla C. Luck, School of Psychology, Curtin University, GPO Box U1987 Perth WA 6845,  
Australia. Email: c.luck@curtin.edu.au

Acknowledgements:

This work was supported by grants number DP180100869 and SR120300015 from the  
Australian Research Council.

Word count: 11166

This is the peer reviewed version of the following article: Luck, C.C. and Patterson, R.R. and Lipp, O.V. 2020. "Prepared" fear or socio-cultural learning? Fear conditioned to guns, snakes, and spiders is eliminated by instructed extinction in a within-participant differential fear conditioning paradigm. *Psychophysiology*. 57 (4): ARTN e13516, which has been published in final form at <https://doi.org/10.1111/psyp.13516>. This article may be used for non-commercial purposes in accordance with Wiley Terms and Conditions for Use of Self-Archived Versions.

### Abstract

Across three experiments, we investigated whether electrodermal responses conditioned to ontogenetic fear-relevant (pointed guns) and phylogenetic fear-relevant stimuli (snakes and spiders) would resist instructed extinction in a within-participant differential fear conditioning paradigm. Instructed extinction involves informing participants before extinction that the unconditional stimulus (US) will no longer be presented. This manipulation has been shown to abolish fear conditioned to fear-irrelevant conditional stimuli, but is said to leave fear conditioned to images of snakes and spiders intact. The latter finding, however, has only been demonstrated when fear-relevance is manipulated between-groups. It is also not known whether instructed extinction affects fear conditioned to ontogenetic fear-relevant stimuli, such as pointed guns. In Experiment 1, we demonstrated that fear conditioned to images of pointed guns does not resist instructed extinction. In Experiment 2, we detected some evidence to suggest that fear conditioned to images of snakes and spiders survives instructed extinction but this evidence was not conclusive. In Experiment 3, we directly compared the effects of instructed extinction on fear conditioned to snakes and spiders and to guns and provide strong evidence that fear conditioned to both classes of stimuli is reduced after instructed extinction with no differences between ontogenetic and phylogenetic stimuli. The current results suggest that when fear relevance is manipulated within participants fear conditioned to both phylogenetic and ontogenetic fear-relevant stimuli responds to instructed extinction providing evidence in favor of a socio-cultural explanation for ‘preparedness’ effects.

Key words: preparedness theory; fear conditioning; instructed extinction; ontogenetic fear-relevant stimuli; guns; phylogenetic fear-relevant stimuli; snakes and spiders

Fear is an adaptive emotional response that facilitates defensive behavior in the presence of environmental threats (Quinn & Fanselow, 2006). All threats, however, may not be equal. Seligman (1970, 1971) challenged the assumption underlying traditional learning theory that any conditional stimulus (CS) could become associated with any unconditional stimulus (US) with equal facility and instead proposed that some aversive associations are hardwired into neural circuitry. Prepared fear associations are said to have developed because they conveyed a survival benefit by facilitating rapid and appropriate responding to threats. Seligman (1970) proposed that prepared associations would show special properties – they would be rapidly acquired, resistant to extinction, and encapsulated from cognitive influence. Furthermore, he proposed that these preferential associations would be established only with aversive outcomes (i.e. selective; would not emerge with an appetitive US). Öhman and Mineka (2001) extended this concept by proposing that prepared fear associations result from dedicated neutral circuitry in the amygdala. This circuitry, known as the ‘fear module’, was argued to respond to phylogenetic, but not ontogenetic, threats automatically and outside of cognitive control.

The concept of prepared fear learning has been the subject of considerable research and debate. While many experiments have reported results consistent with preparedness theory it has also come under considerable challenge (see Davey, 1995, De Jong & Merckelbach 1997; McNally, 1987; Åhs et al., 2018). The classical fear conditioning paradigm has been used to examine whether fear conditioned to different classes of stimuli, meets some, or all, of Seligman’s criteria. During acquisition, a CS (e.g. a fear-relevant image) is paired with an aversive US (e.g. an electric stimulus) and, after repeated pairings, the CS elicits conditional fear responding (e.g. heightened electrodermal responding). During extinction, the CS is presented alone and conditional responding reduces. The criteria of faster acquisition and resistance to

extinction can be examined by comparing the speed of acquisition and extinction between fear-relevant and fear-irrelevant stimuli. The criterion of encapsulation from cognition can be examined using the instructed extinction paradigm (for a review see Luck & Lipp, 2016a). Instructed extinction involves informing participants before extinction that the US will no longer be presented and examining whether this information eliminates conditional responding. Finally, Seligman's criterion of selectivity can be investigated by examining whether the preferential associations present when using aversive USs are absent when non-aversive USs are used.

Phylogenetic animal fear-relevant stimuli (i.e. snakes and spiders) are the most extensively studied class of stimuli and the only class of stimuli, to date, to meet all of Seligman's criteria. Fear conditioned to images of snakes and spiders is rapidly acquired (Ho & Lipp, 2014), resists extinction (Öhman, Erixon, & Lofberg, 1975, Öhman, Erixon, & Olofsson, 1975, Öhman, Fredrickson, Hugdahl, & Rimmö, 1976), and is not affected by instructed extinction (Hugdahl, 1978, Hugdahl & Öhman, 1977, Öhman, Erixon, & Lofberg, 1975, Lipp & Edwards, 2002). Moreover, these preferential associations are only present when using an aversive US (Öhman et al., 1976). Preparedness research has also been extended to social fear-relevant stimuli (i.e. angry faces and other race faces) and to ontogenetic (modern) fear-relevant stimuli (i.e. pointed guns). The findings, however, have been less consistent than for phylogenetic animal fear-relevant stimuli (for a review see Mallan, Lipp, & Cochrane, 2013). Fear conditioned to angry faces and other race faces resists extinction (Öhman & Dimberg, 1978, Dimberg & Öhman, 1983, Olsson, Ebert, Banaji, & Phelps, 2005), but only if the faces used are male (Öhman & Dimberg, 1978, Navarrete et al., 2009). Fear conditioned to images of guns also resists extinction, but only if the guns are pointed towards the participant (Hugdahl & Johnsen, 1989). Unlike fear conditioned to phylogenetic animal fear-relevant stimuli however, conditional

responding to angry faces (Rowles, Lipp, & Mallan, 2012) and other race faces (Mallan, Sax, & Lipp, 2009) is abolished by instructed extinction (i.e. not encapsulated from cognition).

Research, to date, has not examined whether fear conditioned to images of pointed guns would also be removed by instructed extinction, or whether like fear conditioned to images of snakes and spiders, fear conditioned to pointed guns would resist instructed extinction.

There is some research to suggest that fear conditioned to guns could be outside of cognitive control. Conditional fear responding to guns survives a backward masking procedure in which the stimuli are argued to be presented too rapidly to enter into cognitive awareness (Flykt, Esteves, & Öhman, 2007). Furthermore, both ontogenetic and phylogenetic threats have been shown to be preferentially detected in visual search (Blanchette, 2006). Ontogenetic threats, like guns, could only have acquired the ability to rapidly engage with the ‘prepared fear circuitry’ via social and cultural learning, as guns did not pose a survival threat to humans until very recently. Therefore, they should not activate the ‘fear module’ and should not enter into prepared fear associations. In response to evidence of prepared fear learning to guns, Öhman and Mineka (2001) conceded that as the threat posed by a pointed gun is very strong it may be able to gain access to the fear module.

We examined whether fear conditioned to images of pointed guns would respond to instructed extinction in a within-participant differential fear conditioning paradigm. Consistent with the literature on instructed extinction to fear-relevant stimuli, electrodermal responding was used as the primary dependent measure. Participants viewed two images of pointed guns (fear-relevant CS+ and CS-) and two control images of pointed hairdryers (fear-irrelevant CS+ and CS-). During acquisition, the fear-relevant and fear-irrelevant CS+ were followed by a compound US comprising an aversive electric stimulus and a loud noise, while the fear-relevant

and fear-irrelevant CS- were presented alone. After acquisition, participants were informed that the US would no longer be presented and the shock electrode and headphones were removed. We hypothesized that conditional responding to fear-relevant and fear-irrelevant stimuli would be acquired during acquisition, but like fear conditioned to angry faces and other race faces, conditional responding to both guns and hairdryers would be immediately removed by instructed extinction.

## Experiment 1

### Method

**Participants.** Twenty-five (17 female) undergraduate students aged between 17 and 49 years ( $M = 22.96$  years,  $SD = 7.54$ ) volunteered participation in exchange for course credit or monetary compensation of \$AU15. The research protocol was approved by the Curtin University ethics review board. Sample size was based on previous preparedness research which had used a within-participant experimental design (see Lipp, Cronin, Alhadad, & Luck, 2015). Two participants' electrodermal responses during acquisition were lost due to ceiling effects. These participants' data have been included in all other analyses.

**Apparatus/Stimuli.** Conditional stimuli were two pictures of pointed guns (fear-relevant) taken from the international affective picture system (IAPS; image numbers 6230 and 6260; Lang, Bradley, & Cuthbert, 2008) and two pictures of pointed hairdryers<sup>1</sup> (fear-irrelevant). The images were  $1024 \times 768$  pixels in size and were displayed on a 24-inch color LED screen for 6 s. A 200 ms loud white noise blast (95 dB) and electric stimulus combination was used as the unconditional stimulus (US). The electric stimulus was pulsed at 50 Hz and delivered by a

---

<sup>1</sup> Pictures available from the corresponding author on request.

Grass SD9 Stimulator to the participants' preferred forearm. The loud noise had an instantaneous rise time, was generated by a custom built noise generator, and was presented through Sennheiser (HD25-1) headphones. A pseudorandom trial sequence was used, such that a fear-relevant/fear-irrelevant CS+/CS- was not presented more than twice consecutively. Counterbalancing was performed across participants, varying which fear-relevant/fear-irrelevant stimuli were used as the CS+/CS-, whether a fear-relevant or a fear-irrelevant stimulus was presented on the first trial of acquisition, and whether the first trial of acquisition was a CS+/CS-.

A Biopac MP150 system, using AcqKnowledge Version 4.1 at a sampling frequency of 1000 Hz was used to record electrodermal responses and respiration (as a control measure). Electrodermal responding was DC amplified at a gain of 5  $\mu$ Siemens per Volt and recorded using Biopac EL507 electrodes prefilled with an isotonic gel. Respiration was recorded with a respiratory effort transducer attached with an adjustable Velcro strap. DMDX 5.3.4 software (Forster & Forster, 2003) was used to control the stimulus presentation and timing and to record the before and after pleasantness evaluations. Inter-trial interval durations were randomly varied between 11 s, 13 s, and 15 s.

**Procedure.** Participants provided informed consent, washed their hands, and were seated in front of a monitor in a separate room adjacent to the experimenter control room. The respiration belt was attached to the participants' lower torso and the electrodes were placed on the thenar and hypothenar eminences of their non-preferred hand. The shock electrode was placed on the participants' preferred forearm and participants underwent a shock work-up procedure in which the shock intensity was increased from 0 V in steps of 10 V until the participant reported feeling a sensation (usually described as a light tingle). The intensity was then increased in 5 V steps until the participant reported the stimulus to be 'unpleasant, but not

painful' and this intensity was used throughout the experiment. Participants then rated the valence of all four CS pictures with the keyboard on a 1-9 Likert scale (1 = *unpleasant*, 9 = *pleasant*). After the ratings task, participants were informed that they would view pictures on the screen, receive presentations of the electric stimulus, and hear loud blasts of white noise. They were instructed to watch the screen, minimize movement, and pay attention. The conditioning sequence, consisting of habituation, acquisition, and extinction phases, was then started. During habituation, both fear-relevant and fear-irrelevant CSs were presented 4 times each (16 total presentations). During acquisition, both fear-relevant and fear-irrelevant CSs were presented 6 times each, with the offset of the CS+ coinciding with the onset of the US in a 100% reinforcement schedule and the CS- being presented alone. After the last trial of acquisition, the experimenter entered the participants' room and informed them that they would no longer receive the electric stimulus presentations or hear the loud noises, and following the standard instructed extinction procedure, the shock electrode and headphones were removed (see Luck & Lipp, 2015b for a demonstration that shock electrode removal/attachment does not influence instructed extinction effects). After this information, the experimenter left the room and the extinction phase, consisting of 6 presentations of each CS alone (24 presentations total) was started. After the last extinction trial, participants completed a repeat of the ratings task, had the measurement electrodes removed, and were led into the experimenter room to complete the post-experimental questionnaire. The post-experimental questionnaire included an assessment of contingency awareness which required participants to identify which fear-relevant and fear-irrelevant stimuli were paired with the US; separate assessments of white noise and electric stimulus valence on a 7-point Likert scale (-3 = *very unpleasant*, 0 = *neutral*, +3 = *very pleasant*); a post-experimental measure of CS valence on a Likert scale with different anchors (-3

= *very unpleasant*, 0 = *neutral*, +3 = *very pleasant*); and as manipulation check, an assessment of whether the participants believed the instruction that the US would no longer be presented (yes or no question).

**Scoring and Response Definition.** Electrodermal responses were scored in multiple latency windows as recommended by Prokasy and Kumpfer (1973) and Luck and Lipp (2016b). First interval responding was defined as responses starting within 1-4s of CS onset, and second interval responding was defined as responses starting within 4-7s of CS onset. Responses to the US were scored during acquisition as responses starting within 1-4s from US onset (7-10s from CS onset). Both first and second interval responding are sensitive to fear learning, however, first interval responding is also sensitive to orienting and second interval responding is more sensitive to anticipatory processes. The largest response commencing within the defined latency window was scored and the magnitude was calculated as the difference between response onset and response peak (Prokasy & Kumpfer, 1973). Respiration traces were examined as a control measure to identify cases where electrodermal responding was contaminated by deep breaths or excessive moment. No such cases were identified and no responses were excluded.

Electrodermal responses were square root transformed to reduce the positive skew of the distribution (Dawson, Schell, & Filion, 2007) and then range corrected to ensure that each participant was given an even weight in the analyses. The largest response displayed by the participant, usually the response to the first or second US, was used as the reference for the range correction. During habituation, only first interval responses were scored as they are more sensitive to orienting processes (Öhman, 1983) and anticipatory responses would not be expected in this phase.

**Statistical Analyses.** The CS valence evaluations are reported in the supplementary material as they are assessed post-experimentally and therefore not informative regarding the influence of instructed extinction. To confirm that responses to the US during acquisition did not differ between fear-relevant and fear-irrelevant stimuli, the mean scores were subjected to a paired-samples *t*-test. Separate 2 Fear-Relevance (fear-relevant, fear-irrelevant)  $\times$  2 CS (CS+, CS-)  $\times$  *n* Block (habituation = 2, acquisition = 3, extinction = 3) repeated measures ANOVAs were conducted on the first and second interval electrodermal responses for habituation (first interval only), acquisition, and extinction. To examine the influence of instructed extinction separate 2 Fear-Relevance (fear-relevant, fear-irrelevant)  $\times$  2 CS (CS+, CS-)  $\times$  2 Phase (last trial of acquisition, first trial of extinction) repeated measures ANOVAs were conducted on the first and second interval electrodermal responses. We supplemented the instructed extinction follow-up tests with Bayesian analyses to allow evidence for the null hypothesis to be quantified. To be able to provide evidence for/against a difference between fear-relevant and control stimuli we created an instruction score for fear-relevant and fear-irrelevant stimuli separately by subtracting the conditional response (i.e. CS+ - CS-) remaining during the first extinction trial from that present during the last acquisition trial and subjected these scores to a paired sample Bayesian *t*-test. For all Bayesian tests, the effects of interest were compared with paired samples *t*-tests (two-sided; default priors defined in Jasp; Cauchy prior scale: 0.707) and the BF<sub>10</sub> value is reported. According to Jeffreys' (1961) conventions BF<sub>10</sub> values between 0.33 and 1, 0.33 and 0.1, and < 0.1 provide anecdotal, substantial, and strong support for the null hypothesis, respectively. While, BF<sub>10</sub> values between 1 and 3, 3 and 10, and > 10 provide anecdotal, substantial, and strong support for the alternative hypothesis, respectively. It should be noted, however, that Bayes factors should be interpreted as a graded measure of evidence. Multivariate

*F* values (Pillai's Trace) and partial eta-squares are reported for all main effects and interactions. IBM SPSS Statistics 24 was used to conduct all analyses with an alpha level of .05 and JASP 0.9.2 was used to conduct the Bayesian analyses. Analyses were re-run excluding participants who did not pass the contingency assessment (during acquisition, instructed extinction, and extinction) and who reported that they did not believe the instructional manipulation (during instructed extinction and extinction). We report results from the full sample for all analyses but have made note of any instance where exclusion of these participants changes the general pattern of results.

## Results

### Preliminary Analyses

Responses to the US during acquisition did not differ between fear-relevant ( $M = 0.71$ ,  $SD = 0.12$ ) and fear-irrelevant stimuli ( $M = 0.72$ ,  $SD = 0.11$ ),  $t(23) = 0.94$ ,  $p = .359$ . The mean valence ratings for the electric stimuli and noise blasts were -2.04 ( $SD = 0.68$ ) and -1.92 ( $SD = 0.95$ ), respectively. The average US intensity set by the participants was 44.60 volts ( $SD = 19.40$  volts). Six participants were not able to correctly identify the experimental contingencies and three participants reported that they did not believe the instructional manipulation.

### Habituation

**First Interval Responding.** The first interval responses are presented in the upper panel of Figure 1. Fear-relevant stimuli elicited larger responses than fear-irrelevant stimuli,  $F(1, 24) = 17.94$ ,  $p < .001$ ,  $\eta^2 = .428$ . A main effect of block,  $F(1, 24) = 30.40$ ,  $p < .001$ ,  $\eta^2 = .559$ , was moderated by a CS × Block interaction,  $F(1, 24) = 4.29$ ,  $p = .049$ ,  $\eta^2 = .152$ . Unexpectedly, during block one, responses to CS- were marginally higher than to CS+,  $F(1, 24) = 3.99$ ,  $p =$

.057,  $\eta^2 = .143$ , but during block two, responses to CS+ and CS- did not differ,  $F(1, 24) = 2.04$ ,  $p = .166$ ,  $\eta^2 = .078$ . The remaining omnibus effects did not reach significance, all  $F < 1.37$ ,  $p > .254$ ,  $\eta^2 < .054$ .

INSERT FIGURE 1 ABOUT HERE

## Acquisition

**First Interval Responding.** A main effect of CS,  $F(1, 22) = 4.77$ ,  $p = .040$ ,  $\eta^2 = .178$ , a CS  $\times$  Block interaction,  $F(2, 21) = 7.58$ ,  $p = .003$ ,  $\eta^2 = .419$ , and a CS  $\times$  Fear-Relevance interaction,  $F(1, 22) = 5.68$ ,  $p = .026$ ,  $\eta^2 = .205$ , were detected. The CS  $\times$  Block interaction revealed that responses to CS+ and CS- did not differ during blocks one or two, both  $F(1, 22) < 1.97$ ,  $p > .175$ ,  $\eta^2 < .082$ , but during block three,  $F(1, 22) = 17.50$ ,  $p < .001$ ,  $\eta^2 = .443$ , responses to CS+ were larger than to CS-. The CS  $\times$  Fear-Relevance interaction revealed that, across all of acquisition, responding to fear-relevant CS+ was larger than to CS-,  $F(1, 22) = 8.22$ ,  $p = .009$ ,  $\eta^2 = .272$ , but responding between the fear-irrelevant CS+ and CS- did not differ,  $F(1, 22) = 0.23$ ,  $p = .633$ ,  $\eta^2 = .011$ . To confirm that differential responding was present for fear-irrelevant stimuli at the end of acquisition, the analyses were re-run including only the last block of acquisition. These analyses confirmed that responses to CS+ were larger than to CS-, for both fear-relevant,  $F(1, 23) = 15.12$ ,  $p = .001$ ,  $\eta^2 = .397$ , and fear-irrelevant stimuli,  $F(1, 23) = 8.90$ ,  $p = .007$ ,  $\eta^2 = .279$ . The remaining omnibus effects did not reach significance, closest main effects of block,  $F = 3.31$ ,  $p = .056$ ,  $\eta^2 = .240$ , and fear relevance,  $F = 3.81$ ,  $p = .064$ ,  $\eta^2 = .147$ , all other  $F < 0.28$ ,  $p > .759$ ,  $\eta^2 < .026$ .

**Second Interval Responding.** The second interval responding is presented in the bottom panel of Figure 1. A main effect of CS,  $F(1, 22) = 28.66$ ,  $p < .001$ ,  $\eta^2 = .566$ , was moderated by

a CS × Fear-Relevance × Block interaction,  $F(2, 21) = 5.86, p = .009, \eta^2 = .358$ . Responding between fear-relevant CS+ and CS- did not differ during block one,  $F(1, 22) = 1.39, p = .251, \eta^2 = .059$ , but during blocks two and three, responses to fear-relevant CS+ were larger than to fear-relevant CS-, both  $F(1, 22) > 10.52, p < .004, \eta^2 > .323$ . For fear-irrelevant stimuli, responding to CS+ was larger than to CS- during blocks one and three, both  $F(1, 22) > 15.54, p < .001, \eta^2 > .413$ , but did not differ during block two,  $F(1, 22) = 2.14, p = .158, \eta^2 = .089$ . The remaining omnibus effects did not reach significance, closest CS × Block,  $F = 2.67, p = .092, \eta^2 = .203$ , all other  $F < 1.09 p > .355, \eta^2 < .094$ .

### Instructed Extinction

**First Interval Responding.** The first interval responses recorded during the last trial of acquisition and the first trial of extinction are presented in the upper panel of Figure 2. A main effect of CS,  $F(1, 23) = 9.89, p = .005, \eta^2 = .301$ , was moderated by a marginal CS × Phase interaction<sup>2</sup>,  $F(1, 23) = 3.22, p = .086, \eta^2 = .123$ . We performed follow-up analyses due to our a-priori hypotheses and in line with current statistical guidelines cautioning against the use of a rigid  $p$  value cut-off (Wasserstein, 2016), however, the results should be interpreted with care. During the last trial of acquisition, responses to the CS+ were larger than to CS-,  $F(1, 23) = 11.76, p = .002, \eta^2 = .338, BF_{10} = 17.03$ , however, during the first trial of extinction, responding between CS+ and CS- did not differ,  $F(1, 23) = 0.36, p = .553, \eta^2 = .015, BF_{10} = 0.25$ . The  $BF_{10} = 0.85$  for this reduction was indecisive. The instruction effect (i.e. the reduction in differential responding from the last trial of acquisition to the first trial of extinction) did not differ between fear-relevant and fear-irrelevant stimuli,  $BF_{10} = 0.25$ . The remaining omnibus effects did not

---

<sup>2</sup> This interaction does not attain significance in the sample of participants who passed the contingency assessment and manipulation check,  $F(1, 14) = 2.22, p = .158, \eta^2 = .137$ , however, the pattern of follow-up analyses does not change.

reach significance, closest main effect of phase,  $F = 3.04, p = .095, \eta^2 = .117$ , all other,  $F < 1.97, p > .173, \eta^2 < .078$ .

INSERT FIGURE 2 ABOUT HERE

**Second Interval Responding.** The second interval responses recorded during the last trial of acquisition and the first trial of extinction are presented in the bottom panel of Figure 2. Main effects of CS,  $F(1, 22) = 25.16, p < .001, \eta^2 = .533$ , and phase,  $F(1, 22) = 18.65, p < .001, \eta^2 = .459$ , were moderated by a CS  $\times$  Phase interaction,  $F(1, 22) = 17.98, p < .001, \eta^2 = .450$ . During the last trial of acquisition, responses to the CS+ were larger than to CS-,  $F(1, 22) = 42.01, p < .001, \eta^2 = .656, BF_{10} = 11578.12$ , however, conditional responding reduced after instructed extinction,  $BF_{10} = 92.64$ , and during the first trial of extinction, responding between CS+ and CS- did not differ,  $F(1, 22) = 0.71, p = .410, \eta^2 = .031, BF_{10} = 0.29$ . The instruction effect did not differ between fear-relevant and fear-irrelevant stimuli,  $BF_{10} = 0.32$ . The remaining omnibus effects did not reach significance, all  $F < 0.87, p > .361, \eta^2 < .038$ .

## Extinction

**First Interval Responding.** A main effect of block,  $F(2, 23) = 3.90, p = .035, \eta^2 = .253$ , revealed that responding marginally reduced from block one to two,  $p = .080$ , and marginally increased from block two to three,  $p = .060$ . In the sample of participants who could verbalize the experimental contingencies and believed the instructional manipulation, a main effect of CS emerged,  $F(1, 15) = 5.27, p = .036, \eta^2 = .260$ , which revealed that responses to CS+ were larger than to CS-. The remaining omnibus effects did not reach significance, all  $F < 2.91, p > .101, \eta^2 < .108$ .

**Second Interval Responding.** No significant effects were detected, all  $F < 2.08$ ,  $p > .148$ ,  $\eta^2 < .154$ .

## Discussion

In Experiment 1, we examined whether fear conditioned to images of pointed guns would resist instructed extinction after differential fear conditioning. During habituation, images of pointed guns elicited larger electrodermal responses than images of pointed hairdryers. Fear-relevant images are likely more arousing and therefore may elicit higher electrodermal responses. As the habituation phase occurred after the shock work-up procedure, however, it is also possible that this difference occurs due to selective sensitization. Selective sensitization occurs when responses to one class of stimuli are potentiated more than to another class after presentation of an aversive stimulus (see Lovibond, Siddle, & Bond, 1993). Lipp et al. (2015) have demonstrated that electrodermal responding to fear-relevant and fear-irrelevant images of different categories (snakes and spiders vs. fish and birds; angry vs. happy faces, and other race vs. same race faces) does not differ before a shock work-up procedure, but is potentiated to all classes of fear-relevant stimuli after the shock work procedure. In the current study, the shock work-up procedure took place before habituation and therefore it is likely that the enhanced electrodermal responding to images of pointed guns relative to hairdryers also occurs due to selective sensitization. This finding, however, will need to be replicated in a selective sensitization study to show that this difference is absent before the work-up procedure.

Acquisition rates differed between fear-relevant and fear-irrelevant stimuli, however, these differences seemed to be driven by a reduction in the acquisition of first interval electrodermal responding to fear-irrelevant stimuli, rather than the expected faster acquisition to fear-relevant stimuli. Unexpectedly, in the omnibus analysis differential first interval responding

was not present for fear-irrelevant stimuli, but conditioning to these stimuli did emerge by block three of acquisition. This difference was not evident in the second interval electrodermal responses, where differential responding between fear-irrelevant CS+ and CS- was present during blocks one and three of acquisition. It is possible that participants found it more difficult to differentiate between the two hairdryer images, but this possibility is not consistent with the observation that conditioning was present during block one in the second interval responses for fear-irrelevant stimuli. While the cause of the discrepancy is not clear, differential first and second interval responding did emerge for both fear-relevant and fear-irrelevant images by the end of acquisition.

As expected, this differential first and second interval electrodermal responding to both fear-relevant and fear-irrelevant stimuli was abolished immediately by instructed extinction. Furthermore, the Bayesian analyses provided strong evidence against a difference between fear-relevant and fear-irrelevant stimuli in response to instructed extinction. This confirms that, like fear conditioned to angry and other races faces, fear conditioned to images of pointed guns responds to instructed extinction. This appears to be in contrast to studies using images of snakes and spiders – the only class of stimuli, to date, that has been shown to resist instructed extinction.

While there are reports in the literature suggesting that fear conditioned to images of snakes and spiders fulfills all of Seligman's criteria for prepared fear associations, failures to replicate these effects have been reported (for instance see McNally & Foa, 1986; for reviews see McNally, 1987; Åhs et al., 2018) suggesting that prepared fear associations obtained to images of snakes and spiders may not be robust. There is also a large procedural discrepancy between instructed extinction studies examining social and ontogenetic fear-relevant stimuli and those examining phylogenetic fear-relevant stimuli as all demonstrations of resistance to instructed

extinction of fear conditioned to snakes and spiders have manipulated fear-relevance between-participants. It is not clear whether the same result would emerge in a within subject design.

Although seemingly minor, this discrepancy could be critical (see Mertens, Raes, & De Houwer, 2016 for a discussion). In a between-participants design, one group of participants is exposed to fear-relevant images, while another is exposed to fear-irrelevant images. Repeated presentations of fear-relevant images could increase participants' state anxiety levels and therefore enhance conditioning. This would not occur in the group exposed to fear-irrelevant images, leading to a potential group difference in fear learning. This problem does not occur in the within-participants design because participants are exposed to both fear-relevant and fear-irrelevant images, and therefore any increases in state anxiety caused by the fear-relevant images will also enhance conditioning to the fear-irrelevant images. As a group difference in state anxiety could explain the different response to instructed extinction between social and ontogenetic fear-relevant stimuli and phylogenetic fear-relevant stimuli, it is necessary to replicate resistance to instructed extinction of fear conditioned to images of snakes and spiders using a within-participants design.

## **Experiment 2**

To confirm that resistance to instructed extinction for snakes and spiders reported in the literature occurs due to an intrinsic property of snakes and spiders rather than a procedural difference in how fear relevance is manipulated, we conducted a replication of Experiment 1 using images of snakes and spiders as the fear-relevant stimuli and images of fish and birds as the fear-irrelevant stimuli. We hypothesized that conditional responding to images of snakes and spiders would remain intact after instructed extinction, but conditional fear responses to images of fish and birds would be immediately abolished by instructed extinction.

## **Method**

**Participants.** Twenty-five (18 female) undergraduate students aged between 18 and 26 years ( $M = 21.44$ ,  $SD = 2.36$ ) volunteered participation in exchange for course credit or monetary compensation of \$AU15. The research protocol was approved by the Curtin University ethics review board. All participants provided complete datasets.

**Apparatus/Stimuli.** Pictures of a snake and spider were used as fear-relevant CSs and pictures of a fish and a bird were used as fear-irrelevant stimuli<sup>3</sup>. All other aspects of the procedure were identical to Experiment 1.

## Results

### Preliminary Analyses

Responses to the US during acquisition did not differ between fear-relevant ( $M = 0.67$ ,  $SD = 0.08$ ) and fear-irrelevant stimuli ( $M = 0.67$ ,  $SD = 0.09$ ),  $t(24) = 0.04$ ,  $p = .972$ . The mean valence ratings for the electric stimuli and noise blasts were -1.64 ( $SD = 1.04$ ) and -1.96 ( $SD = 1.10$ ), respectively. The average US intensity set by the participants was 53.08 volts ( $SD = 19.31$  volts). Three participants were not able to identify the experimental contingencies and one participant reported that they did not believe the instructional manipulation.

### Habituation

**First Interval Responding.** The first interval responding is presented in the upper panel of Figure 3. Responding to fear-relevant stimuli was larger than to fear-irrelevant stimuli,  $F(1, 24) = 19.91$ ,  $p < .001$ ,  $\eta^2 = .453$ , and responding during block one was larger than during block

---

<sup>3</sup> Pictures available from the corresponding author on request

two,  $F(1, 24) = 10.70, p = .003, \eta^2 = .308$ . The remaining omnibus effects did not reach significance, all  $F < 1.57, p > .223, \eta^2 < .062$ .

INSERT FIGURE 3 ABOUT HERE

## Acquisition

**First Interval Responding.** Responding to CS+ was larger than to CS-,  $F(1, 24) = 31.02, p < .001, \eta^2 = .564$ , and responding to fear-relevant stimuli was larger than to fear-irrelevant stimuli,  $F(1, 24) = 5.78, p = .024, \eta^2 = .194$ . The remaining omnibus effects did not reach significance, all  $F < 2.36, p > .117, \eta^2 < .170$ .

**Second Interval Responding.** The second interval responding is presented in the bottom panel of Figure 3. A main effect of CS,  $F(1, 24) = 37.12, p < .001, \eta^2 = .607$ , was moderated by a CS  $\times$  Block interaction,  $F(2, 23) = 4.67, p = .020, \eta^2 = .289$ . Responses to CS+ were larger than to CS- during all blocks of acquisition,  $F(1, 24) > 12.34, p < .002, \eta^2 > .339$ , but increased in magnitude from blocks 1 to 3,  $p = .005$ . The remaining omnibus effects did not reach significance, closest CS  $\times$  Fear Relevance,  $F = 3.19, p = .087, \eta^2 = .117$ , CS  $\times$  Fear-Relevance  $\times$  Block<sup>4</sup>,  $F = 2.80, p = .081, \eta^2 = .196$ , all other  $F < 0.39, p > .541, \eta^2 < .016$ .

## Instructed Extinction

**First Interval Responding.** The first interval responses recorded during the last trial of acquisition and the first trial of extinction are presented in the upper panel of Figure 4. Main effects of CS,  $F(1, 24) = 8.57, p = .007, \eta^2 = .263$ , and phase,  $F(1, 24) = 22.73, p < .001, \eta^2 =$

---

<sup>4</sup>This interaction reflects that during block one, differential responding was not present for fear-relevant stimuli,  $F(1, 24) = 1.87, p = .184, \eta^2 = .072$ , but was for fear-irrelevant stimuli,  $F(1, 24) = 11.34, p = .003, \eta^2 = .321$ ,

.486, were moderated by a marginal CS × Fear Relevance × Phase interaction<sup>5</sup>,  $F(1, 24) = 3.63, p = .069, \eta^2 = .131$ . We performed follow-up analyses due to our a-priori hypotheses and in line with current statistical guidelines cautioning against using a rigid  $p$  value cut-off (Wasserstein, 2016), however, the results should be considered exploratory and interpreted with care. During the last trial of acquisition, responses to the CS+ were larger than to CS- for fear-irrelevant stimuli,  $F(1, 24) = 6.31, p = .019, \eta^2 = .208, \text{BF}_{10} = 2.78$ , but unexpectedly this difference was not significant for fear-relevant stimuli,  $F(1, 24) = 2.32, p = .141, \eta^2 = .088, \text{BF}_{10} = 0.58$ . During the first trial of extinction, responses to CS+ were larger than to CS- for fear-relevant,  $F(1, 24) = 6.20, p = .020, \eta^2 = .205, \text{BF}_{10} = 2.68$ , but not for fear-irrelevant stimuli,  $F(1, 24) = 0.86, p = .363, \eta^2 = .035, \text{BF}_{10} = 0.31$ . A comparison across phase revealed that differential responding did not change for fear-relevant stimuli,  $F(1, 24) = 0.05, p = .821, \eta^2 = .002, \text{BF}_{10} = 0.22$ , but significantly reduced for fear-irrelevant stimuli,  $F(1, 24) = 4.76, p = .039, \eta^2 = .165, \text{BF}_{10} = 1.56$ . A Bayesian comparison of the size of the instruction effect (i.e. the reduction in differential responding from the last trial of acquisition to the first trial of extinction) between fear-relevant and fear-irrelevant stimuli was indecisive,  $\text{BF}_{10} = 1.00$ . The remaining omnibus effects did not reach significance, all  $F < 2.50, p > .127, \eta^2 < .095$ .

#### INSERT FIGURE 4 ABOUT HERE

**Second Interval Responding.** The second interval responses recorded during the last trial of acquisition and the first trial of extinction are presented in the bottom panel Figure 4.

---

<sup>5</sup> This interaction is not significant in the sample of participants who could verbalize the experimental contingencies and reported believing the instructional manipulation,  $F(1, 20) = 0.98, p = .334, \eta^2 = .047$ , instead a CS × Phase interaction,  $F(1, 20) = 8.48, p = .009, \eta^2 = .298$ , revealed that differential responding was present during the last trial of acquisition,  $F(1, 20) = 18.20, p < .001, \eta^2 = .476, \text{BF}_{10} = 85.04$ , and during the first trial of extinction,  $F(1, 20) = 4.73, p = .042, \eta^2 = .191, \text{BF}_{10} = 1.57$ , but reduced in magnitude after the instruction,  $\text{BF}_{10} = 5.72$ .

Main effects of CS,  $F(1, 24) = 13.24, p = .001, \eta^2 = .355$ , and phase,  $F(1, 24) = 8.51, p = .008, \eta^2 = .262$ , were moderated by a CS  $\times$  Phase,  $F(1, 24) = 16.51, p < .001, \eta^2 = .408$ , and a CS  $\times$  Fear Relevance  $\times$  Phase interaction,  $F(1, 24) = 9.18, p = .006, \eta^2 = .277$ . During the last trial of acquisition, responses to the CS+ were larger than to CS- for fear-relevant,  $F(1, 24) = 6.28, p = .019, \eta^2 = .207$ ,  $BF_{10} = 2.76$ , and for fear-irrelevant stimuli,  $F(1, 24) = 34.91, p < .001, \eta^2 = .593, BF_{10} = 4706.90$ . During the first trial of extinction, however, responding did not differ between CS+ and CS- for fear-relevant,  $F(1, 24) = 1.32, p = .262, \eta^2 = .052, BF_{10} = 0.38$ , or for fear-irrelevant stimuli,  $F(1, 24) = 1.33, p = .261, \eta^2 = .052, BF_{10} = 0.38$ . A comparison across phase revealed that differential responding did not change for fear-relevant stimuli,  $F(1, 24) = 1.26, p = .273, \eta^2 = .050, BF_{10} = 0.37$ , but significantly reduced for fear-irrelevant stimuli,  $F(1, 24) = 31.04, p < .001, \eta^2 = .564, BF_{10} = 2189.93$ . A Bayesian comparison of the size of the instruction effect revealed that the magnitude of the reduction in differential responding was larger for fear-irrelevant than for fear-relevant stimuli,  $BF_{10} = 7.61$ . Further-exploration, however, revealed that this difference in the effect of instructions was likely driven by a larger conditional response during the last trial of acquisition for fear-irrelevant stimuli in comparison to fear-relevant stimuli,  $F(1, 24) = 5.24, p = .031, \eta^2 = .179, BF_{10} = 1.87$ , and therefore, this result needs to be interpreted with care as it likely does not reflect greater resistance to instructed extinction for fear-relevant stimuli. The remaining omnibus effects did not reach significance, all  $F < 1.20, p > .284, \eta^2 < .048$ .

## Extinction

**First Interval Responding.** A marginal main effect of block,  $F(2, 23) = 3.38, p = .052, \eta^2 = .227$ , a marginal CS  $\times$  Fear-Relevance interaction,  $F(1, 24) = 3.22, p = .085, \eta^2 = .118$ , and a marginal CS  $\times$  Block interaction,  $F(2, 23) = 2.72, p = .087, \eta^2 = .191$ , were detected. We

performed follow-up analyses but they should be considered exploratory and interpreted with care. The marginal CS × Fear-Relevance interaction revealed that, across extinction, responding to the fear-relevant CS+ was larger than to the fear-relevant CS-,  $F(1, 24) = 6.20, p = .020, \eta^2 = .205$ , but responding did not differ between fear-irrelevant CS+ and CS-,  $F(1, 24) = 0.04, p = .838, \eta^2 = .002$ . The marginal CS × Block interaction revealed that, responses to CS+ were larger than to CS- during block one,  $F(1, 24) = 4.92, p = .036, \eta^2 = .170$ , and two (marginal),  $F(1, 24) = 3.60, p = .070, \eta^2 = .130$ , but did not differ during block three,  $F(1, 24) = 0.60, p = .448, \eta^2 = .024$ . The remaining omnibus effects did not reach significance, all  $F < 2.79, p > .108, \eta^2 < .105$ .

**Second Interval Responding.** No omnibus effects were detected, closest Fear Relevance × Block<sup>6</sup>, all  $F = 2.99, p = .070, \eta^2 = .206$ , all other,  $F < 1.07, p > .361, \eta^2 < .085$ .

## Discussion

In Experiment 2, we examined whether fear conditioned to images of snakes and spiders would resist instructed extinction in a within-participant differential conditioning design. This was done to confirm that previous reports of resistance to instructed extinction were not due to stronger conditioning caused by increased levels of state-anxiety in the group exposed to fear-relevant images relative to controls. Conditional first and second interval electrodermal responding was acquired during acquisition and this development was not faster for fear-relevant stimuli. There was some evidence of a differential influence of instructed extinction between fear-relevant and fear-irrelevant stimuli – however, overall the evidence was weak and inconclusive. After instructed extinction, conditional first interval responding reduced for fear-

<sup>6</sup> The marginal Fear Relevance × Block interaction reflects that during block two responding to fear-irrelevant stimuli was larger than to fear-relevant stimuli,  $F(1, 24) = 5.70, p = .025, \eta^2 = .192$ .

irrelevant stimuli but not for fear-relevant stimuli, resulting in an elimination of differential first interval responding on the first trial of extinction for fear-irrelevant stimuli, but not for fear-relevant stimuli. This pattern may suggest that fear conditioned to snakes/spiders survives instructed extinction but needs to be interpreted with care as there was no significant differential first interval responding during the last trial of acquisition for fear-relevant stimuli and the Bayesian comparison between fear-relevant and fear-irrelevant stimuli was inconclusive. In the second interval responding, conditional responding was present at the end of acquisition and eliminated by the first trial of extinction for both classes of stimuli. Interestingly, however, while the magnitude of conditional responding reduced after instructed extinction for fear-irrelevant stimuli, it did not for fear-relevant stimuli and the Bayesian analyses provided substantial evidence that the instruction effect was larger for fear-irrelevant stimuli than on fear-relevant stimuli. These differences, however, likely *do not* provide evidence that fear conditioned to snakes and spiders resists instructed extinction. The magnitude of the conditional second interval response was significantly larger during the last trial of acquisition for fear-irrelevant stimuli and therefore the larger reduction that occurs for fear-irrelevant stimuli likely reflects that a larger conditional response was present to respond to the instruction for this class of stimuli. While these findings do provide some evidence that fear conditioned to snakes and spiders resists instructed extinction, due to the problems outlined above the results need to be replicated before any firm conclusions can be drawn. Furthermore, to provide strong evidence for a difference between ontogenetic and phylogenetic stimuli in response to instructed extinction the two classes of stimuli should be compared in a single within-participant study.

### Experiment 3

In Experiment 3, we assessed the effect of instructed extinction on fear conditioned to ontogenetic and phylogenetic fear-relevant stimuli within a single study. Comparing classes of fear-relevant stimuli in a within-participant design ensures that any differences between phylogenetic and ontogenetic stimuli are due to the characteristics of the stimuli themselves and not due to factors that can vary between studies or between groups. We used the same within-participant fear conditioning design as in Experiments 1 and 2, however, phylogenetic fear-relevant (snakes/spiders) and ontogenetic fear-relevant stimuli (guns) were used as the two classes of stimuli. Based on the results of Experiments 1 and 2, we hypothesized conditional responding to images of snakes and spiders would remain intact after instructed extinction, but conditional responding to images of guns would be immediately abolished after instructed extinction.

## Method

**Participants.** Eighty-four (67 female; 15 male; 1 other) undergraduate students aged between 17 and 45 years ( $M = 20.80$ ,  $SD = 4.31$ ) volunteered participation in exchange for course credit. The research protocol was approved by the Curtin University ethics review board. Three participants were excluded from all measures (one reported that they could not feel the electric stimulus; one was run through acquisition twice due to experimenter error and one left during the experiment), electrodermal data of three additional participants were not scored as visual inspection of the data revealed that they did not show orienting or responding to the USs, and one participant was excluded from the ratings data as they indicated after the experiment that they had misunderstood the scale, leaving a final sample of 78 and 80 for electrodermal responses and ratings, respectively. An a priori power calculated revealed that 74 participants would be required to be 80% confident of detecting a significant CS × Fear Relevance × Phase

interaction. The power analysis was conducted using G\*Power 3.1.9.4 and based on the effect size for the CS × Fear Relevance × Phase interaction from the first interval responses in Experiment 2 (i.e.  $\eta^2 = .131$ ;  $f = 0.388$ ). Power was calculated based on an ANOVA  $F$  test (repeated measures within factors; 1 group; 4 measurement points [phylogenetic and ontogenetic during the last trial of acquisition; phylogenetic and ontogenetic during the first extinction trial]; Effect size specification: as in Cohen (1988); nonsphericity correction: 1) using the CS difference scores (CS+ - CS-). Difference scores were used for the power calculations as a main effect of phase detected using the CS difference scores is statistically equivalent to a CS × Phase interaction and a Fear Relevance × Phase interaction detected using the CS difference scores is statistically equivalent to a CS × Fear Relevance × Phase interaction.

**Apparatus/Stimuli.** The fear-relevant pictures used in Experiments 1 and 2 were used as the ontogenetic (guns) and phylogenetic fear-relevant CSs (snake and spider), respectively. All other aspects of the procedure were identical to Experiment 1.

## Results

### Preliminary Analyses

Responses to the US during acquisition did not differ between phylogenetic ( $M = 0.74$ ,  $SD = 0.15$ ) and ontogenetic stimuli ( $M = 0.73$ ,  $SD = 0.15$ ),  $t(77) = 0.74$ ,  $p = .462$ . The mean valence ratings for the electric stimuli and noise blasts were -1.94 ( $SD = 0.73$ ) and -2.07 ( $SD = 0.88$ ), respectively. The average US intensity set by the participants was 45.74 volts ( $SD = 17.20$  volts). Twenty-one participants were not able to identify the experimental contingencies and thirteen participants reported that they did not believe the instructional manipulation.

### Habituation

**First Interval Responding.** The first interval responding is presented in the upper panel of Figure 5. Responses declined from block one to two,  $F(1, 77) = 42.91, p < .001, \eta^2 = .358$ . The remaining omnibus effects did not reach significance, all  $F < 1.31, p > .256, \eta^2 < .017$ .

## Acquisition

**First Interval Responding.** Main effects of CS,  $F(1, 77) = 74.70, p < .001, \eta^2 = .492$ , and block,  $F(2, 76) = 28.02, p < .001, \eta^2 = .424$ , and CS  $\times$  Fear Relevance,  $F(1, 77) = 5.00, p = .028, \eta^2 = .061$ , and CS  $\times$  Block,  $F(2, 76) = 8.25, p = .001, \eta^2 = .178$ , interactions were detected. The CS  $\times$  Block interaction revealed that responses were higher to CS+ than to CS- during all blocks, all  $F(1, 77) > 15.10, p < .001, \eta^2 > .164$ , but differential responding increased in magnitude from blocks 1 to 3,  $p = .001$ . Follow up of the CS  $\times$  Fear Relevance interaction revealed that conditional responding was present to both phylogenetic,  $F(1, 77) = 66.03, p < .001, \eta^2 = .462$ , and ontogenetic stimuli,  $F(1, 77) = 36.31, p < .001, \eta^2 = .320$ , but the magnitude of this differential responding was larger for phylogenetic stimuli. This difference was driven by heightened responses to the ontogenetic CS- in comparison with the phylogenetic CS-,  $F(1, 77) = 6.61, p = .012, \eta^2 = .079$ . Responses to the phylogenetic and ontogenetic CS+ did not differ,  $F(1, 77) = 0.05, p = .833, \eta^2 = .001$ . The remaining omnibus effects did not reach significance, all  $F < 2.63, p > .108, \eta^2 < .033$ .

INSERT FIGURE 5 ABOUT HERE

**Second Interval Responding.** The second interval responding is presented in the bottom panel of Figure 5. A main effect of CS,  $F(1, 77) = 36.96, p < .001, \eta^2 = .324$ , and a CS  $\times$  Block interaction,  $F(2, 76) = 8.24, p = .001, \eta^2 = .178$ , were moderated by a CS  $\times$  Fear Relevance  $\times$  Block interaction  $F(2, 76) = 3.74, p = .028, \eta^2 = .090$ . For phylogenetic stimuli, conditional

responding was present during all blocks, all  $F(1, 77) > 10.68, p < .002, \eta^2 > .121$ , however, for ontogenetic stimuli, conditional responding was not present during block one,  $F(1, 77) = 0.46, p = .500, \eta^2 = .006$ , but was during blocks two and three, both  $F(1, 77) > 5.94, p < .018, \eta^2 > .071$ . The remaining omnibus effects did not reach significance, closest CS  $\times$  Fear Relevance,  $F = 2.91, p = .092, \eta^2 = .036$ , all other  $F < 1.68, p > .193, \eta^2 < .043$ .

### Instructed Extinction

**First Interval Responding.** The first interval responses recorded during the last trial of acquisition and the first trial of extinction are presented in the upper panel of Figure 6. Main effects of CS,  $F(1, 77) = 35.49, p < .001, \eta^2 = .315$ , and phase,  $F(1, 77) = 75.47, p < .001, \eta^2 = .495$ , were moderated by a CS  $\times$  Phase interaction,  $F(1, 77) = 7.05, p = .010, \eta^2 = .084$ . Responses to CS+ were larger than to CS- during the last trial of acquisition,  $F(1, 77) = 30.10, p < .001, \eta^2 = .281$ ,  $BF_{10} = 28172.27$ , and during the first trial of extinction,  $F(1, 77) = 8.78, p = .004, \eta^2 = .102$ ,  $BF_{10} = 6.97$ , but the magnitude of the conditional response decreased after instructed extinction,  $BF_{10} = 3.27$ . A comparison of the instruction score (i.e. the reduction in differential responding from the last trial of acquisition to the first trial of extinction) between ontogenetic and phylogenetic stimuli provided evidence against a difference between them,  $BF_{10} = 0.13$ . In the sample of participants who passed the contingency assessment and manipulation check, the CS  $\times$  Phase interaction,  $F(1, 49) = 13.75, p = .001, \eta^2 = .219$ , reveals that although responses were larger to CS+ than to CS- during the last trial of acquisition,  $F(1, 49) = 31.70, p < .001, \eta^2 = .393$ ,  $BF_{10} = 18795.74$ , the magnitude of the conditional response decreased after instructed extinction,  $BF_{10} = 50.18$ , and responses to CS+ and CS- did not differ during the first trial of extinction,  $F(1, 49) = 2.54, p = .118, \eta^2 = .049$ ,  $BF_{10} = 0.50$ . In this sample, the comparison of the instruction score between ontogenetic and phylogenetic stimuli also provided

evidence against a difference between them,  $BF_{10} = 0.20$ . The remaining omnibus effects did not reach significance, all  $F < 1.01$ ,  $p > .318$ ,  $\eta^2 < .013$ .

INSERT FIGURE 6 ABOUT HERE

**Second Interval Responding.** The second interval responses recorded during the last trial of acquisition and the first trial of extinction are presented in the bottom panel of Figure 6. Main effects of CS,  $F(1, 77) = 18.73$ ,  $p < .001$ ,  $\eta^2 = .196$ , and phase,  $F(1, 77) = 33.10$ ,  $p < .001$ ,  $\eta^2 = .301$ , were moderated by a CS  $\times$  Phase interaction,  $F(1, 77) = 8.10$ ,  $p = .006$ ,  $\eta^2 = .095$ . Responses to CS+ were larger than to CS- during the last trial of acquisition,  $F(1, 77) = 18.09$ ,  $p < .001$ ,  $\eta^2 = .190$ ,  $BF_{10} = 328.11$ , but the magnitude of the conditional response reduced after instructed extinction,  $BF_{10} = 5.17$ , and responding between CS+ and CS- did not differ on the first trial of extinction,  $F(1, 77) = 1.66$ ,  $p = .202$ ,  $\eta^2 = .021$ ,  $BF_{10} = 0.28$ . A comparison of the instruction score between ontogenetic and phylogenetic stimuli provided evidence against a difference between them,  $BF_{10} = 0.13$ . The remaining omnibus effects did not reach significance, all  $F < 2.29$ ,  $p > .134$ ,  $\eta^2 < .029$ .

## Extinction

**First Interval Responding.** A main effect of CS,  $F(1, 77) = 6.95$ ,  $p = .010$ ,  $\eta^2 = .083$ , was moderated by a CS  $\times$  Block interaction,  $F(2, 76) = 4.52$ ,  $p = .014$ ,  $\eta^2 = .106$ , which revealed that responses were higher to CS+ than to CS- during block one,  $F(1, 77) = 14.59$ ,  $p < .001$ ,  $\eta^2 = .159$ , but did not differ during blocks two or three, both  $F(1, 77) < 0.02$ ,  $p > .895$ ,  $\eta^2 < .001$ . In the sample of participants who passed the contingency assessment and could verbalize the experimental contingencies the CS  $\times$  Block interaction does not attain significance,  $F(2, 48) =$

$1.70, p = .195, \eta^2 = .066$ . The remaining omnibus effects did not reach significance, all  $F < 1.55, p > .219, \eta^2 < .040$ .

**Second Interval Responding.** No significant effects were detected, all  $F < 1.96, p > .148, \eta^2 < .049$ .

## Discussion

In Experiment 3, we examined whether instructed extinction would have differential effects on fear conditioned to phylogenetic and ontogenetic fear-relevant stimuli when compared within the same study. No differences between the two classes of stimuli were detected in habituation and differential first and second interval responding were successfully acquired during acquisition to both classes of stimuli. The magnitude of the conditional first interval responding was larger for phylogenetic fear-relevant stimuli but this was driven by increased responding to the ontogenetic CS- compared to the phylogenetic CS-. For second interval responding, conditioning emerged one block earlier for phylogenetic fear-relevant stimuli which may suggest that acquisition is faster for phylogenetic fear-relevant stimuli.

Unexpectedly, fear conditioned to both phylogenetic and ontogenetic fear-relevant stimuli showed a clear reduction after instructed extinction and we observed substantial evidence against a difference between them. In the full sample of participants, conditional first interval responding was evident during the last trial of acquisition and during the first trial of extinction but was significantly reduced in size after instructed extinction. This reduction was not moderated by the class of stimuli. In the sample of participants who could verbalize the contingencies and who passed the manipulation check, however, conditional responding was not present during the first trial of extinction, likely suggesting that the remnant detected in the full

sample was driven by participants who did not believe the manipulation or were not aware of the contingencies. In the second interval responses, the pattern of results was very similar – conditional responding was present at the end of acquisition but was immediately eliminated after instructed extinction with strong support against a difference between ontogenetic and phylogenetic fear-relevant stimuli being obtained in the Bayesian analyses. These results suggest that fear conditioned to *both* ontogenetic and phylogenetic fear-relevant stimuli reduces after instructed extinction and provide evidence that neither are ‘encapsulated from cognition’.

### General Discussion

The current investigation examined whether fear conditioned to images of pointed guns (ontogenetic fear-relevant stimuli) and fear conditioned to images of snakes and spiders (phylogenetic fear-relevant stimuli) would resist instructed extinction. This study provides the first demonstration that fear conditioned to images of guns responds to instructed extinction and is the first study to examine whether reports of resistance to instructed extinction for fear conditioned to snakes and spiders would replicate when fear-relevance is manipulated within-participants. In Experiment 1, after differential conditioning with images of pointed hairdryers and pointed guns, conditional electrodermal first and second interval responding was immediately eliminated by instructed extinction. In Experiment 2, participants underwent differential acquisition to phylogenetic fear-relevant (snakes and spiders) and fear-irrelevant (fish and birds) animals. After instructed extinction, there was some evidence that differential first interval responses to fear-relevant animals stayed intact but as differential first interval responding was not present for fear-relevant stimuli during the last trial of acquisition the pattern of results was inconclusive. Similarly, differential second interval responding was eliminated for fear-relevant and fear-irrelevant stimuli, but the instruction effect was larger for fear-irrelevant

stimuli – however, this likely reflects that differential second interval responding was larger for this class of stimuli during the last trial of acquisition. We conducted Experiment 3, to confirm the pattern of results for phylogenetic fear-relevant stimuli (snakes and spiders) and to directly compare this class of stimuli with ontogenetic fear-relevant stimuli (pointed guns). Conditional first and second interval responding to both phylogenetic and ontogenetic fear-relevant stimuli was immediately reduced by instructed extinction and the Bayesian analyses provided substantial evidence against a difference between them. Although conditional second interval responding was abolished on the first trial of extinction, conditional first interval responding, although significantly reduced, remained evident for both ontogenetic and phylogenetic fear-relevant stimuli. The significant reduction in conditional responding observed after instructed extinction provides strong evidence against the proposal that fear conditioned to snakes and spiders is ‘encapsulated from cognition’ but the remnant detected after the manipulation could indicate some resistance to the instruction for both classes of fear-relevant stimuli. Unfortunately, as conditioning to ontogenetic and phylogenetic fear-relevant stimuli was directly compared in this experiment, no comparison to fear-irrelevant stimuli is available. It is very well established in the literature, however, that fear conditioned to fear-irrelevant stimuli is immediately abolished by instructed extinction (for a review see Luck & Lipp, 2016a) and this pattern was detected for fear-irrelevant stimuli in Experiments 1 and 2 as well. In the sample of participants who reported that they believed the instructional manipulation and who could verbalize the experimental contingencies, however, conditional responding was *not present* on the first trial of extinction. Therefore it seems that the remnant of differential conditioning was likely driven by the subset of participants who did not believe the instructions.

On the whole, the current investigation provides strong evidence that fear conditioned to ontogenetic and phylogenetic fear-relevant stimuli is not ‘encapsulated from cognition’ as the magnitude of the conditional response reduced significantly after instructed extinction and conditional responding was eliminated in participants who passed the manipulation check and could identify the experimental contingencies. Resistance to instructed extinction for fear conditioned to snakes and spiders has been reported in the literature but the evidence is mixed. Of the seven published experiments to examine whether instructed extinction eliminates fear conditioned to snakes and spiders five (Öhman, Erixon, and Löfberg, 1975; Hugdahl & Öhman, 1977; Hugdahl, 1978; Soares & Öhman, 1993; Lipp & Edwards, 2002) provide evidence for resistance to the instruction, whereas two did not (Cook, Hodes, & Lang, 1986; Experiments 4 and 6). The current study, however, is the first to examine whether fear conditioned to snakes and spiders resists instructed extinction in a within-participant design and thus eliminates the possibility that differences in state anxiety levels between fear-relevance groups could drive the differential effects of instruction. The within-participant comparison also provides more power than the between-participant comparisons and the sample size used in Experiment 3 was larger than those of all other past investigations which should provide a more reliable estimation of the population effect.

The influence of instructed extinction on fear conditioned to snakes and spiders and to guns detected in Experiment 3 is in line with reports that fear conditioned to angry faces (see Rowles et al., 2012) and to other race faces (see Mallan et al., 2009) is reduced by instructed extinction. Preparedness theory is not unchallenged (see Davey, 1995; De Jong & Merckelbach 1997; McNally, 1987; Åhs et al., 2018) and in our opinion, the strongest evidence in support of preparedness theory was the finding that fear conditioned to snakes and spiders resists instructed

extinction. The current findings are an important addition to the literature as they suggest that when the more reliable within-participant design is used fear conditioned to snakes and spiders, like that conditioned to all other classes of fear-relevant stimuli, is subject to verbal instruction. This evidence against the Seligman's 'encapsulation from cognition' criterion suggests that fear conditioned to snakes and spiders does not respond differently to fear conditioned to social or ontogenetic fear-relevant stimuli. Thus, it seems more parsimonious, to us at least, to conclude that fear associations with phylogenetic, social, and ontogenetic fear-relevant stimuli are not different and therefore evidence for facilitated fear learning to fear-relevant stimuli (relative to fear-irrelevant stimuli) is likely due to socio-cultural learning. Thus the current investigation provides evidence against prepared fear learning – it seems that snakes and spiders are not so 'special' after all.

### References

- Åhs, F., Rosén, J., Kastrati, G., Fredrikson, M., Agren, T., & Lundström, J. N. (2018). Biological preparedness and resistance to extinction of skin conductance responses conditioned to fear relevant animal pictures: A systematic review. *Neuroscience & Biobehavioral Reviews*, 95, 430-437. <https://doi.org/10.1016/j.neubiorev.2018.10.017>
- Blanchette, I. (2006). Snakes, spiders, guns, and syringes: How specific are evolutionary constraints on the detection of threatening stimuli? *The Quarterly Journal of Experimental Psychology*, 59(8), 1484-1504.  
<https://doi.org/10.1080/02724980543000204>
- Cook, E. W., Hodes, R. L., & Lang, P. J. (1986). Preparedness and phobia: Effects of stimulus content on human visceral conditioning. *Journal of Abnormal Psychology*, 95(3), 195-207. <https://doi.org/10.1037/0021-843X.95.3.195>
- Davey, G. C. L. (1995). Preparedness and phobias: Specific evolved associations or a generalized expectancy bias? *Behavioral and Brain Sciences*, 18(2), 289-297.  
<https://doi.org/10.1017/S0140525X00038498>
- Dawson, M. E., Schell, A. M., & Filion, D. L. (2007). The electrodermal system. In J. T. Cacioppo, L.G. Tassinary & G.G. Bernston (Eds.), (2007). *Handbook of Psychophysiology* (pp. 159-181). Cambridge: Cambridge University Press.
- de Jong, P. J., & Merckelbach, H. (1997). No convincing evidence for a biological preparedness explanation of phobias. *Behavioral and Brain Sciences*, 20(2), 362-363.  
<https://doi.org/10.1017/S0140525X97221461>

- Dimberg, U., & Öhman, A. (1983). The Effects of Directional Facial Cues on Electrodermal Conditioning to Facial Stimuli. *Psychophysiology*, 20(2), 160-167.  
<https://doi.org/10.1111/j.1469-8986.1983.tb03282.x>
- Flykt, A., Esteves, F., & Öhman, A. (2007). Skin conductance responses to masked conditioned stimuli: Phylogenetic/ontogenetic factors versus direction of threat? *Biological psychology*, 74(3), 328-336. <https://doi.org/10.1016/j.biopsych.2006.08.004>
- Forster, K., & Forster, J. (2003). DMDX: A Windows display program with millisecond accuracy. *Behavior Research Methods, Instruments, & Computers*, 35, 116-124.  
<https://doi.org/10.3758/BF03195503>
- Ho, Y., & Lipp, O. V. (2014). Faster acquisition of conditioned fear to fear-relevant than to nonfear-relevant conditional stimuli. *Psychophysiology*, 51(8), 810-813.  
<https://doi.org/10.1111/psyp.12223>
- Hugdahl, K., & Johnsen, B. H. (1989). Preparedness and electrodermal fear-conditioning: Ontogenetic vs phylogenetic explanations. *Behaviour Research and Therapy*, 27(3), 269-278. [https://doi.org/10.1016/0005-7967\(89\)90046-6](https://doi.org/10.1016/0005-7967(89)90046-6)
- Hugdahl, K., & Öhman, A. (1977). Effects of instruction on acquisition and extinction of electrodermal responses to fear-relevant stimuli. *Journal of Experimental Psychology: Human Learning and Memory*, 3(5), 608-618. <https://doi.org/10.1037/0278-7393.3.5.608>
- Hugdahl, K. (1978). Electrodermal conditioning to potentially phobic stimuli: Effects of instructed extinction. *Behaviour Research and Therapy*, 16(5), 315-321.  
[https://doi.org/10.1016/0005-7967\(78\)90001-3](https://doi.org/10.1016/0005-7967(78)90001-3)
- Jeffreys, H. (1961). *Theory of Probability* (3 ed.). Oxford, UK: Oxford University Press.

Lipp, O. V., & Edwards, M. S. (2002). Effect of instructed extinction on verbal and autonomic indices of Pavlovian learning with fear-relevant and fear-irrelevant conditional stimuli.

*Journal of Psychophysiology*, 16(3), 176-186. <https://doi.org/10.1027//0269-8803.16.3.176>

Lipp, O. V., Cronin, S. L., Alhadad, S. S. J., & Luck, C. C. (2015). Enhanced sensitization to animal, interpersonal, and intergroup fear-relevant stimuli (but no evidence for selective one-trial fear learning). *Psychophysiology*, 52(11), 1520-1528.  
<https://doi.org/10.1111/psyp.12513>

Lovibond, P. F., Siddle, D. A. T., & Bond, N. W. (1993). Resistance to extinction of fear-relevant stimuli: Preparedness or selective sensitization? *Journal of Experimental Psychology: General*, 122, 449-461. <https://doi.org/10.1037/0096-3445.122.4.449>

Luck, C. C., & Lipp, O. V. (2016a). Instructed extinction in human fear conditioning: History, recent developments, and future directions. *Australian Journal of Psychology*, 68(3), 209-227. <https://doi.org/10.1111/ajpy.12135>

Luck, C. C., & Lipp, O. V. (2015b). To remove or not to remove? Removal of the unconditional stimulus electrode does not mediate instructed extinction effects. *Psychophysiology*, 52(9), 1248-1256. <https://doi.org/10.1111/psyp.12452>

Luck, C. C., & Lipp, O. V. (2016a). Instructed extinction in human fear conditioning: History, recent developments, and future directions. *Australian Journal of Psychology*, 68(3), 209-227. <https://doi.org/10.1111/ajpy.12135>

Luck, C. C., & Lipp, O. V. (2016b). When orienting and anticipation dissociate — a case for scoring electrodermal responses in multiple latency windows in studies of human fear

conditioning. *International Journal of Psychophysiology*, 100, 36-43.

<https://doi.org/10.1016/j.ijpsycho.2015.12.003>

Mallan, K. M., Lipp, O. V., & Cochrane, B. (2013). Slithering snakes, angry men and out-group members: What and whom are we evolved to fear? *Cognition and Emotion*, 27(7), 1168-1180. <https://doi.org/10.1080/02699931.2013.778195>

Mallan, K. M., Sax, J., & Lipp, O. V. (2009). Verbal instruction abolishes fear conditioned to racial out-group faces. *Journal of Experimental Social Psychology*, 45(6), 1303-1307. <https://doi.org/10.1016/j.jesp.2009.08.001>

Mertens, G., Raes, A. K., & De Houwer, J. (2016). Can prepared fear conditioning result from verbal instructions? *Learning and Motivation*, 53, 7-23.

<https://doi.org/10.1016/j.lmot.2015.11.001>

Navarrete, C. D., Olsson, A., Ho, A. K., Mendes, W. B., Thomsen, L., & Sidanius, J. (2009). Fear Extinction to an Out-Group Face. *Psychological Science*, 20(2), 155-158. <https://doi.org/10.1111/j.1467-9280.2009.02273.x>

McNally, R. J., & Foa, E. B. (1986). Preparedness and resistance to extinction to fear-relevant stimuli: A failure to replicate. *Behaviour Research and Therapy*, 24(5), 529-535. [https://doi.org/10.1016/0005-7967\(86\)90033-1](https://doi.org/10.1016/0005-7967(86)90033-1)

McNally, R. J. (1987). Preparedness and phobias: A review. *Psychological Bulletin*, 101(2), 283-303. <https://doi.org/10.1037/0033-2909.101.2.283>

O'Brien, F., & Cousineau, D. (2014). Representing Error bars in within-subject designs in typical software packages. *Tutorials in Quantitative Methods for Psychology*, 10, 56-67. <https://doi.org/10.20982/tqmp.10.1.p056>

Öhman, A., & Dimberg, U. (1978). Facial expressions as conditioned stimuli for electrodermal responses: A case of "preparedness"? *Journal of Personality and Social Psychology*, 36(11), 1251-1258. <https://doi.org/10.1037/0022-3514.36.11.1251>

Öhman, A., & Mineka, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. *Psychological review*, 108(3), 483-522.  
<https://doi.org/10.1037/0033-295X.108.3.483>

Öhman, A. (1983). The orienting response during Pavlovian conditioning. In D. A. T. Siddle (Ed.), *Orienting and habituation: Perspectives in human research* (pp. 315-370). New York: Wiley.

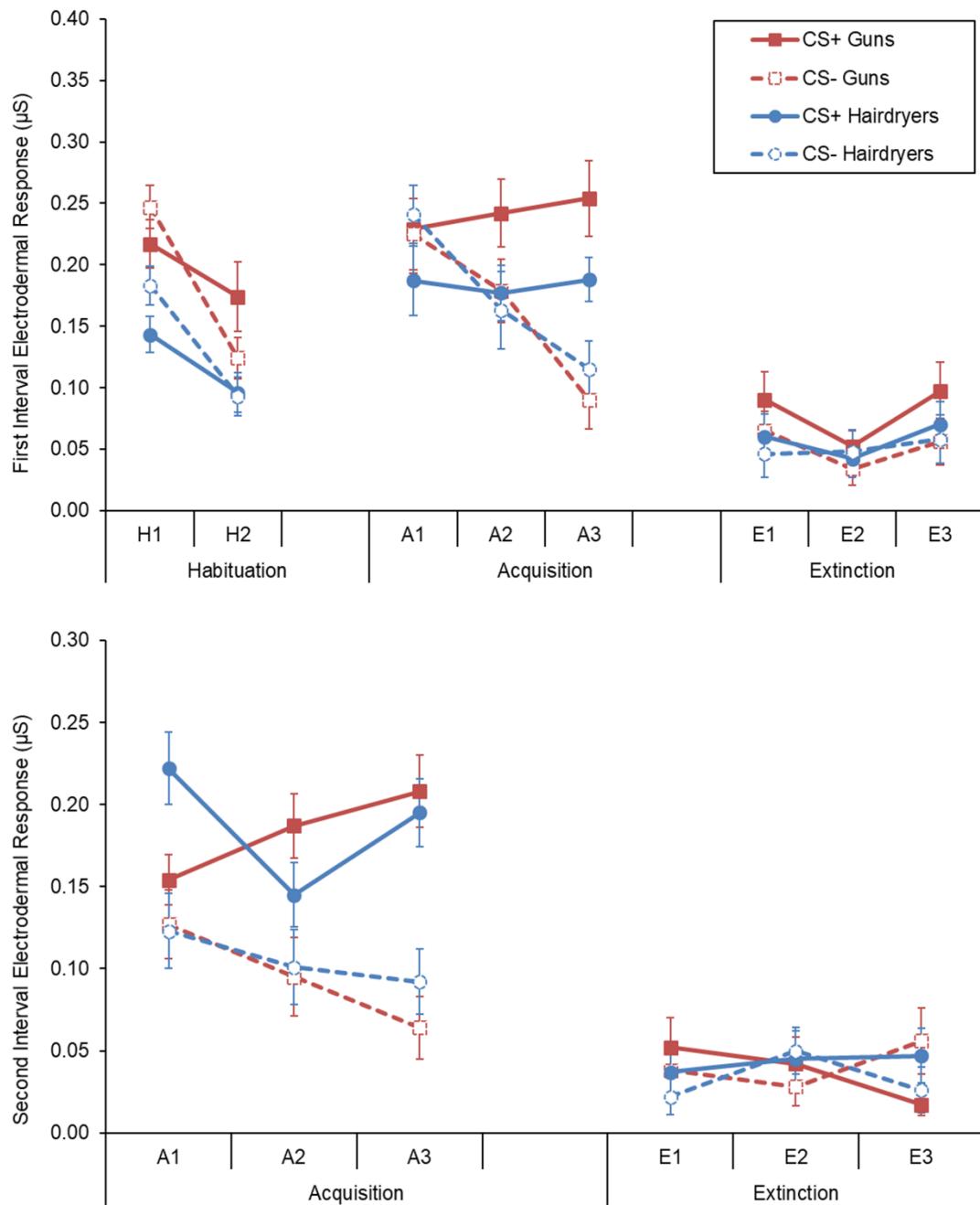
Öhman, A., Erixon, G., & Löfberg, I. (1975). Phobias and preparedness: Phobic versus neutral pictures as conditioned stimuli for human autonomic responses. *Journal of Abnormal Psychology*, 84(1), 41-45. <https://doi.org/10.1037/h0076255>

Öhman, A., Eriksson, A., & Olofsson, C. (1975). One-Trial Learning and Superior Resistance to Extinction of Autonomic Responses Conditioned to Potentially Phobic Stimuli. *Journal of comparative and physiological psychology*, 88(2), 619-627.  
<https://doi.org/10.1037/h0078388>

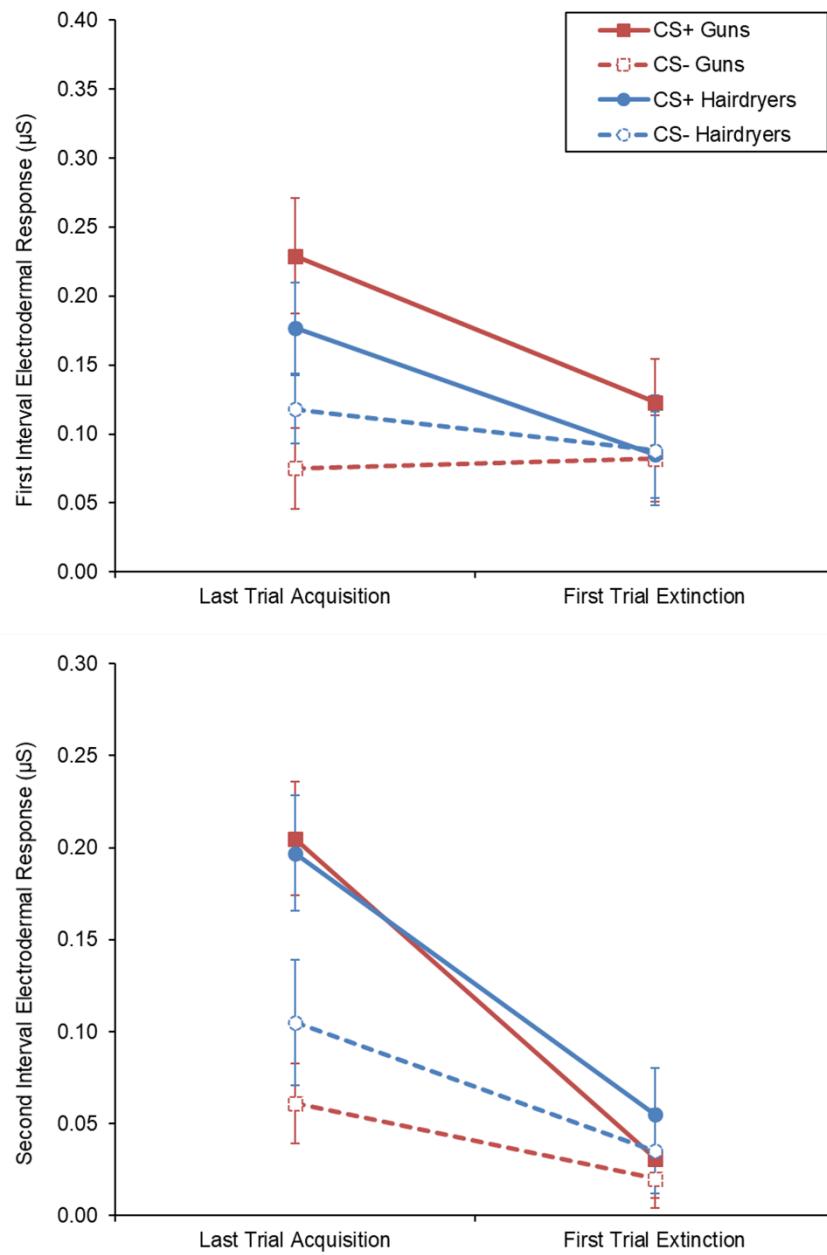
Öhman, A., Fredrikson, M., Hugdahl, K., & Rimmö, P.-A. (1976). The premise of equipotentiality in human classical conditioning: Conditioned electrodermal responses to potentially phobic stimuli. *Journal of Experimental Psychology: General*, 105(4), 313-337. <https://doi.org/10.1037/0096-3445.105.4.313>

Olsson, A., Ebert, J. P., Banaji, M. R., & Phelps, E. A. (2005). The Role of Social Groups in the Persistence of Learned Fear. *Science*, 309(5735), 785-787.  
<https://doi.org/10.1126/science.1113551>

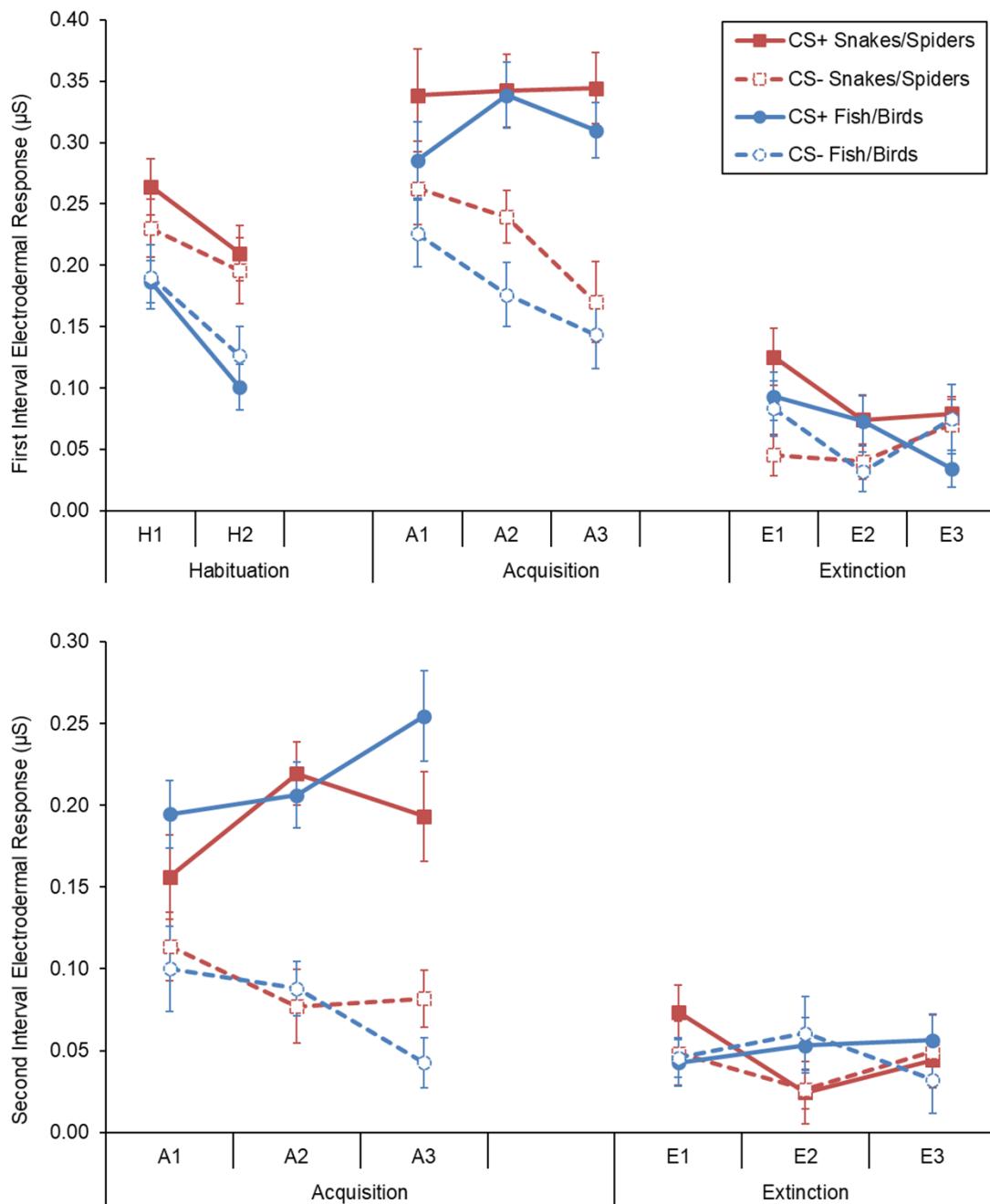
- Prokasy, W. F., & Kumpfer, K. L. (1973). Classical Conditioning. In W. F. Prokasy & D. C. Raskin (Eds.), *Electrodermal Activity in Psychological Research* (pp. 157-202). U.S.A: Academic Press.
- Quinn, J. J., & Fanselow, M. S. (2006). Defenses and memories: Functional neural circuitry of fear and conditional responding. In M. G. Craske, D. Hermans & D. Vansteenwegen (Eds.), (2006). Fear and learning: From basic processes to clinical implications (pp. 37-52). Washington: APA Books.
- Rowles, M. E., Lipp, O. V., & Mallan, K. M. (2012). On the resistance to extinction of fear conditioned to angry faces. *Psychophysiology*, 49(3), 375-380.  
<https://doi.org/10.1111/j.1469-8986.2011.01308.x>
- Seligman, M. E. (1970). On the generality of the laws of learning. *Psychological review*, 77(5), 406-418. <https://doi.org/10.1037/h0029790>
- Seligman, M. E. P. (1971). Phobias and preparedness. *Behavior therapy*, 2(3), 307-320.  
[https://doi.org/10.1016/S0005-7894\(71\)80064-3](https://doi.org/10.1016/S0005-7894(71)80064-3)
- Soares, J. J. F., & Öhman, A. (1993). Preattentive processing, preparedness and phobias: Effects of instruction on conditioned electrodermal responses to masked and non-masked fear-relevant. *Behaviour Research and Therapy*, 31(1), 87-95. [https://doi.org/10.1016/0005-7967\(93\)90046-W](https://doi.org/10.1016/0005-7967(93)90046-W)
- Wasserstein, R. L., & Lazar, N. A. (2016). The ASA Statement on p-Values: Context, Process, and Purpose. *The American Statistician*, 70(2), 129-133.  
<https://doi.org/10.1080/00031305.2016.1154108>



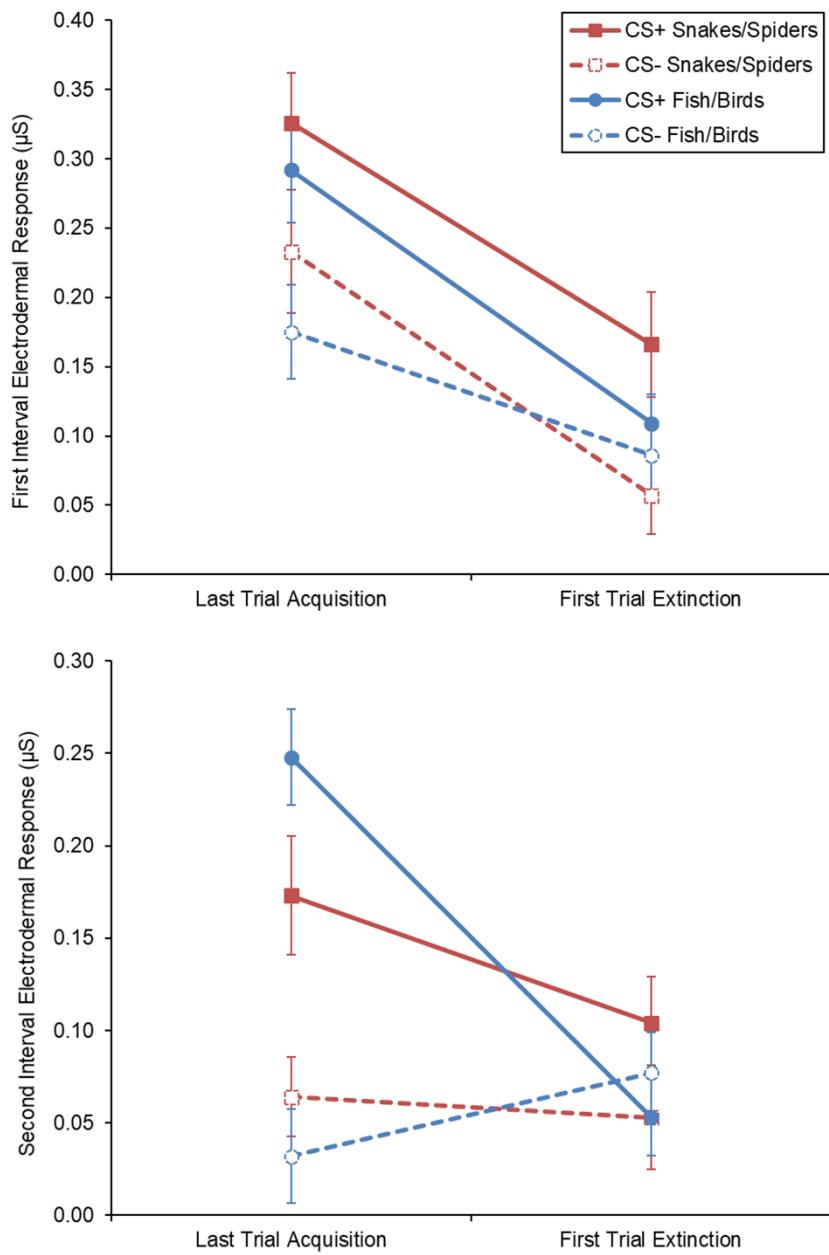
*Figure 1.* First (top panel) and second (bottom panel) interval electrodermal responding recorded to fear-relevant (guns) and fear-irrelevant (hairdryers) stimuli during Experiment 1. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).



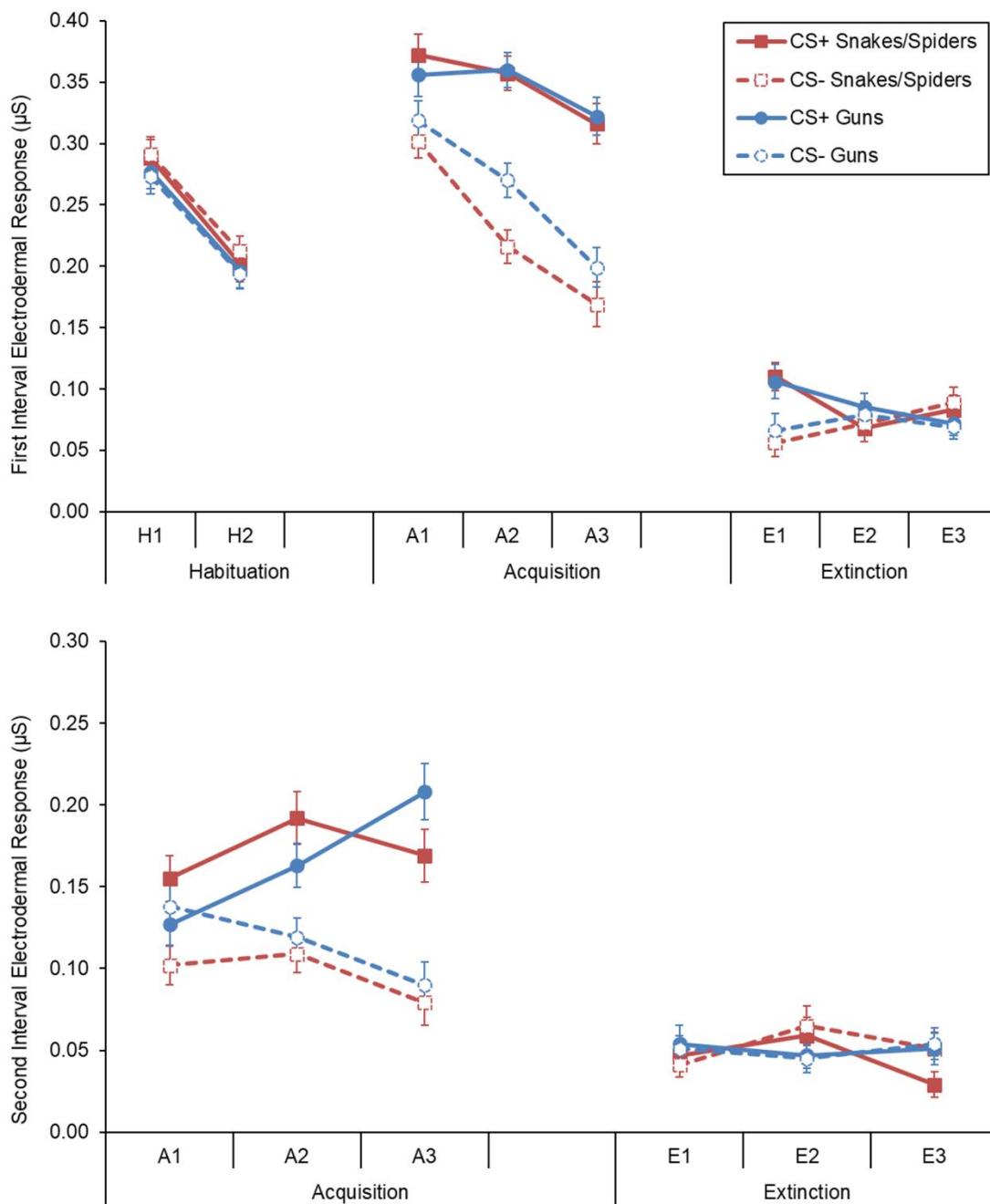
*Figure 2.* First (top panel) and second (bottom panel) interval electrodermal responding recorded to fear-relevant (guns) and fear-irrelevant (hairdryers) stimuli during the last trial of acquisition and immediately after instructed extinction (first trial of extinction) during Experiment 1. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).



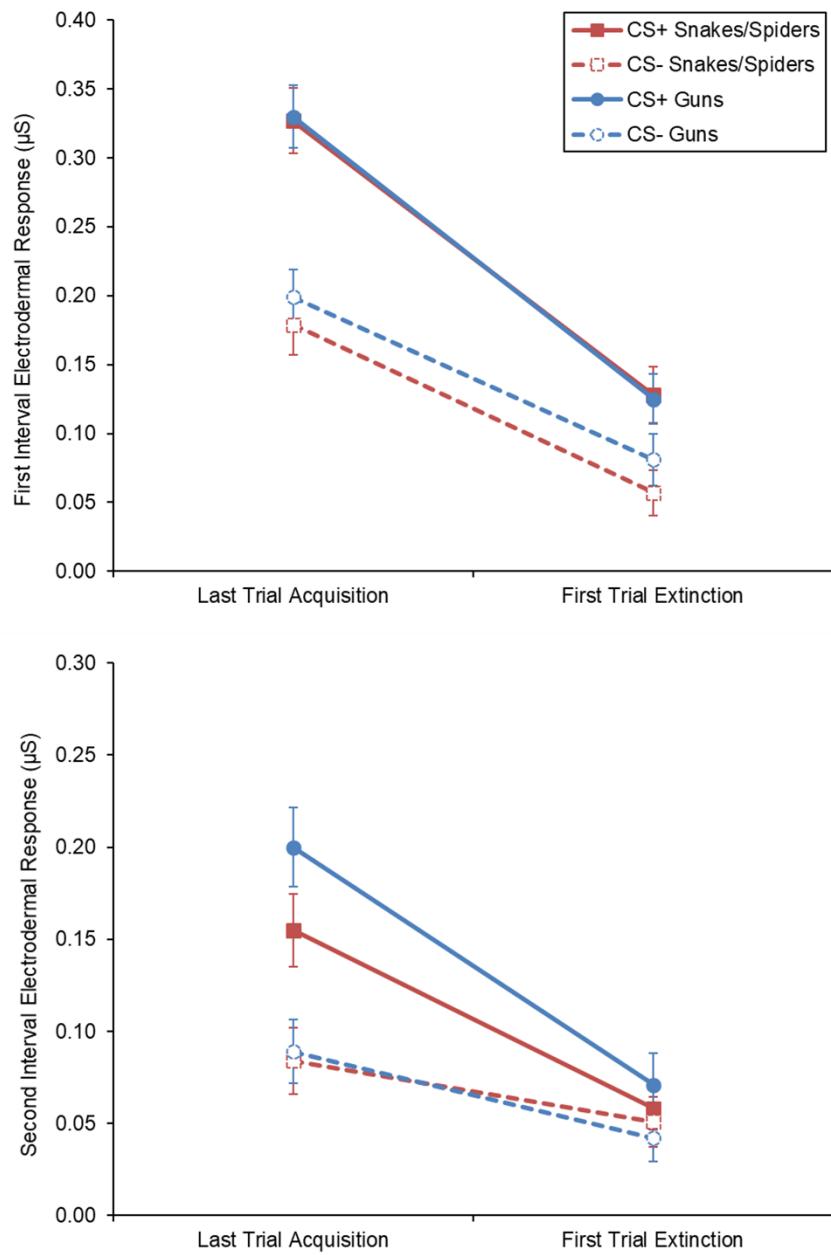
*Figure 3.* First (top panel) and second (bottom panel) interval electrodermal responding recorded to fear-relevant (snakes/spiders) and fear-irrelevant (fish/birds) stimuli during Experiment 2. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).



*Figure 4.* First (top panel) and second (bottom panel) interval electrodermal responding recorded to fear-relevant (snakes/spiders) and fear-irrelevant (fish/birds) stimuli during the last trial of acquisition and immediately after instructed extinction (first trial of extinction) during Experiment 2. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).



*Figure 5.* First (top panel) and second (bottom panel) interval electrodermal responding recorded to phylogenetic fear-relevant (snakes/spiders) and ontogenetic fear-relevant (guns) stimuli during Experiment 3. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).



*Figure 6.* First (top panel) and second (bottom panel) interval electrodermal responding recorded to phylogenetic fear-relevant (snakes/spiders) and ontogenetic fear-relevant (guns) stimuli during the last trial of acquisition and immediately after instructed extinction (first trial of extinction) during Experiment 3. Error bars represent SEMs for within-participant designs based on O'Brien and Cousineau (2014).