Received: 3 February 2020 | Revised: 24 June 2020 | Accepted: 7 July 2020

DOI: 10.1111/psyp.13650

ORIGINAL ARTICLE

PSYCHOPHYSIOLOGY SPR WILEY

Inconsistent analytic strategies reduce robustness in fear extinction via skin conductance response

Luke John Ney[1] | **Patrick A.F. Laing**[2] | **Trevor Steward**[3] | **Daniel V. Zuj**[4] | **Simon Dymond**[4,5] | **Kim L. Felmingham**[3]

[1]School of Psychology, University of Tasmania, Hobart, TAS, Australia

[2]Melbourne Neuropsychiatry Centre, Department of Psychiatry, University of Melbourne & Melbourne Health, Carlton, VIC, Australia

[3]School of Psychological Sciences, University of Melbourne, Carlton, VIC, Australia

[4]Department of Psychology, Swansea University, Swansea, UK

[5]Department of Psychology, Reykjavik University, Reykjavik, Iceland

Correspondence
Luke John Ney, School of Medicine (Psychology), University of Tasmania, Private Bag 30, Sandy Bay, TAS 7005, Australia.
Email: luke.ney@utas.edu.au

Funding information
National Health and Medical Research Council, Grant/Award Number: APP1073041

Abstract

Robustness of fear conditioning and extinction paradigms has become increasingly important for many researchers interested in improving the study of anxiety and trauma disorders. We recently illustrated the wide variability in data analysis techniques in this paradigm, which we argued may result in a lack of robustness. In the current study, we resampled data from six of our own fear acquisition and extinction data sets, with skin conductance as the outcome. In the resampled and original data sets, we found that effect sizes that were calculated using discrepant statistical strategies, sourced from a non-exhaustive search of high-impact articles, were often poorly correlated. The main contributors to poor correlations were the selection of trials from different stages of each experimental phase and the use of average compared to trial-by-trial analysis. These findings reinforce the importance of focusing on robustness in the psychophysiological measurement of fear acquisition and extinction in the laboratory and may guide prospective researchers in which decisions may most impact the robustness of their results.

KEYWORDS

fear conditioning, fear extinction, robustness, skin conductance response, statistical analysis, threat conditioning

1 | INTRODUCTION

Anxiety disorders are characterized by excessive and persistent aversive responses to neutral, safe, or ambiguous stimuli (Craske et al., 2009; Grupe & Nitschke, 2013). Similarly, deficient learning and retention of fear extinction has been proposed as a primary maintaining factor in anxiety and post-traumatic stress (PTSD) disorders (Graham, Callaghan, & Richardson, 2014; Grupe & Nitschke, 2013; Suarez-Jimenez et al., 2019; Zuj & Norrholm, 2019; Zuj, Palmer, Lommen, & Felmingham, 2016). Improved understanding of the underlying mechanisms of extinction could aid the development of clinical interventions for anxiety and traumatic disorders (Craske, Treanor, Conway, Zbozinek, & Vervliet, 2014; Lebois, Seligowski, Wolff, Hill, & Ressler, 2019).

Recent decades have seen increasingly sophisticated measurements of fear acquisition and extinction in the laboratory, with important implications for the treatment of anxiety and PTSD (Milad & Quirk, 2012; Zuj & Norrholm, 2019). Fear acquisition paradigms model adaptive threat learning via contingent pairings of previously neutral conditioned stimuli (CS) and innately aversive unconditioned stimuli (US). Fear (or threat) extinction procedures feature repeated unreinforced presentation of the conditioned threat stimulus (CS+), leading to decreased threat responses and new safety learning that competes with previous threat memories (Bouton, 2004). Extinction learning and the subsequent retention of the extinction memory can be quantified by comparing the extinguished CS+ and the CS− during the extinction and retention phases, respectively. Responses during the extinction phase

can be used to index extinction learning itself, while differences at subsequent testing reflect retention or consolidation of extinction or retained fear memories (Lonsdorf et al., 2017; Milad & Quirk, 2012).

Phasic skin conductance responses (SCRs) constitute the most commonly used measure of conditioned threat responding (Bach et al., 2018; Lonsdorf et al., 2017; Pittig, Treanor, LeBeau, & Craske, 2018). The amplitude of physiological responding to a threat signal (i.e., the CS+) can be compared to the safety signal (i.e., the CS−) to infer extinction. Physiological measures—especially SCRs—are notoriously noisy, with large degrees of individual variance and biological artifacts (Bach et al., 2018; Boucsein, 2012; Ojala & Bach, 2020). We had previously expressed concerns that, due to insufficient power in most studies, slight variations on core analytical strategies—such as choice of statistical analysis or removal of trials—might result in inconsistent findings in the same paradigm (Ney et al., 2018). The high-impact studies that we surveyed in this publication differed in the number and order of trials included in the analysis, in which trials were averaged, and whether differential responding was used. Previously, high heterogeneity in experimental design and analysis of studies examining reinstatement effects following extinction was reported (Haaker, Golkar, Hermans, & Lonsdorf, 2014). More recently, Lonsdorf, Merz, and Fullana (2019) expressed concern that no consensus currently exists among fear extinction studies estimating the extinction retention index, which was originally developed as a way of inferring retention of extinction memory relative to responding during acquisition. Lonsdorf et al. identified 16 separate analysis strategies and showed that these strategies, despite claiming to be measuring a single underlying construct (i.e., extinction retention), were, in fact, partly poorly correlated and did not necessarily reflect extinction memory.

Research domains that are generally underpowered, have flexible outcomes, and are evaluated using multiple analytical strategies are at high risk of poor replicability (Ioannidis, 2005; Simmons, Nelson, & Simonsohn, 2011). In the present study, we sought to examine the similarity of results produced by variations in statistical analyses of fear acquisition and extinction. Doing so was intended as an extension from Lonsdorf, Klingelhöfer-Jens, et al. (2019) where only the robustness of the extinction retention index was tested. Our aim was to test the robustness of the analytical strategies for analyzing SCRs during acquisition and extinction learning from high-impact studies. To do this, we performed a non-exhaustive literature search to gather several contrasting statistical strategies for similar fear conditioning paradigms. We then correlated the effect sizes of different methods obtained across multiple of our own data sets, which we resampled to create a final sample of $N = 40$ data sets. We hypothesized that slight variations of analytical strategies would result in weak, non-significant correlational effect sizes, despite the methods purportedly measuring the same constructs.

2 | METHOD

2.1 | Method selection

We searched online data sets (PubMed, PsycINFO, Web of Science) for keywords "fear acquisition," "fear conditioning," "fear extinction," "skin conductance," and "extinction." To ensure that we obtained a sufficiently influential yet not overwhelmingly large sample, articles that had 150 or more citations on Web of Science and were published post-2000 were included in the first-pass search. Due to data sets from our lab consisting of within-session CS± differential acquisition paradigms with SCRs as the primary outcome measure, there were several restrictions on the studies that were included. First, we did not include studies that had used contextual or additional CS+ manipulations during fear acquisition or extinction learning. Second, we only included analyses from day 1 of multi-day paradigms, so long as they included both fear acquisition and extinction learning phases in a single session. Finally, only studies using SCRs as a primary outcome measure were included, since SCRs are the predominant acquisition measure and there has been significant heterogeneity in its scoring and reporting. Strategies were separated into three categories. Some studies had focused on the difference between SCRs from the acquisition to extinction phase (ACQ − EXT), whereas others were either interested in the change in SCRs over the extinction phase ($EXT_{early} - EXT_{late}$) or in estimating a gross measure of fear extinction learning during that phase (EXT). We were aware of several other articles with fewer than 150 citations that had used unique analysis strategies; these were added to increase the pool of strategies for the ACQ − EXT and $EXT_{early} - EXT_{late}$ methods (see Table 1).

2.2 | Data sets

We used data from six of our own data sets for this analysis (details below). To increase the sample size, data were resampled with replacement from the six data sets to create an additional 34 data sets of $N = 60$ each. Resampling was performed using the Resampling Stats Add-in for Excel v4.0 (Simon, Bruce, & Troiana, 2013). Resampling with replacement was preferred to ensure higher variability of the resampled data sets to the original data sets. To ensure that resampled data sets would mimic interphase correlations of SCRs, we resampled by row; that is, each resample consisted

TABLE 1 Description of different strategies for measuring extinction learning using skin conductance responses

Analytic strategy	Strategy #	# of trials	Trials included	Trial analysis	Stimuli analysis	Analysis	Study
ACQ − EXT	Strategy 1	8 (ACQ), 16 (EXT)	All (ACQ), last 2 (EXT)	Average	Diff	Phase × Group	Graham and Milad (2013)
	Strategy 2	5 (ACQ), 10 (EXT)	Maximum Response (ACQ), Last 2 (EXT)	Average	Diff	Phase × Group	Milad et al. (2010)
	Strategy 3	8 (ACQ), 7 (EXT)	All (ACQ), last 3 (EXT)	Average	Diff	Phase × Group	White and Graham (2016)[c]
	Strategy 4	20 (ACQ), 20 (EXT)	Last half (ACQ), First half (EXT)	Average, using paired t-test contrasts[a]	Diff	Phase × Group	Grady et al. (2016)[c]
EXT	Strategy 1	16	Last three-quarters	Average	CS+, CS−	Group × Stim	Milad et al. (2009)
	Strategy 2	5	All	Trial-by-trial	CS+, CS−	Trial × Group × Stim	Zuj et al. (2016)[c]
	Strategy 3	16	Last half	Average	CS+, CS−	Group × Stim	Garfinkel et al. (2014)
	Strategy 4	10	Last trial	One trial	Diff	Group	Schiller et al. (2010)
	Strategy 5	10	Last 2	Average	CS+, CS−	Group × Stim	Milad et al. (2008)
	Strategy 6	5	All	Running average[b]	Diff	Trial × Group	Milad et al. (2006)
	Strategy 7	8	First 2	Trial-by-trial	Diff	Trial × Group	Pace-schott et al. (2013)[c]
EXT_{early} − EXT_{late}	Strategy 1	6	First half, second half	Average	CS+, CS−	Phase × Group × Stim	Blechert et al. (2007)
	Strategy 2	14	First half, second half	Average	Diff	Phase × Group	Michael et al. (2007); Phelps et al. (2004)
	Strategy 3	16	First quarter, last quarter	Average	CS+	Phase × Group	Milad et al. (2013)
	Strategy 4	32, 16	First half, second half	Average	CS+	Phase × Group	Soliman et al. (2010); Zeidan et al. (2011)
	Strategy 5	10	All	Linear contrast	CS+, CS−	Trial × Group × Stim	Lovibond et al. (2009)[c]; Ney et al. (in prep)[c]

Abbreviations: ACQ, acquisition; CS+, conditioned stimulus to the aversive unconditioned stimulus; CS−, Conditioned stimulus as a safety signal; Stim, stimulus type (CS+ vs. CS−); Diff, differential; EXT, extinction.

[a]This study was the only study to use a test other than ANOVA.

[b]Running average response was calculated with trials one and two averaged as a single response, trials two and three averaged, and so on.

[c]Study was not identified as part of the original search criteria but was added post hoc due to methodological variance.

of the entire phase of one participant's CS+ or CS− response (but not both). This ensured that the data would mimic real responding as closely as possible without resampling any participant's entire differential response.

All data sets used either red and blue (data sets 1, 2, and 3) or green and orange (data sets 4, 5, and 6) circles as CS, presented on a computer screen. In all studies, CS+ and CS− were randomized between participants. CS duration was 12s with intertrial intervals of 12–21s ($M = 16$ s). Each study consisted of three phases: habituation, acquisition, and extinction learning. Habituation lasted for four trials (i.e., four separate presentations of CS+ and four of CS−) and the extinction phase consisted of 10 trials. Data sets 1–3 featured five acquisition trials, while data sets 4–6 had 7. For the latter data sets, only the first five trials were analyzed, so to be consistent with data sets 1–3 during analysis and resampling. Although data sets 3–6 were 2-day paradigms, only the first day was used to be consistent with data sets 1 and 2. Data sets 1–3 had a 100% CS − US reinforcement schedule during acquisition, whereas the other data sets had a 62.5% schedule.

Each of the original data sets had a different group manipulation. For data sets 1–3 ($N = 120$, $N = 56$ and $N = 79$, respectively) participants consisted of PTSD-diagnosed cases, trauma-exposed cases and non-trauma exposed cases (each data set had a different manipulation outside of this, see publications or Supporting Information for additional details; Hsu et al. in prep; Ney et al. in prep; Zuj, Palmer, Hsu, et al., 2016). In data set 4, the group manipulation was sham or anodal transcranial direct current stimulation (tDCS) to the dorsal lateral prefrontal cortex prior to or following the extinction learning phase ($N = 80$, Ney et al., in prep; Vicario et al., 2020). In data set 5, the group manipulation was naturally cycling women in the early follicular phase of the menstrual cycle compared to women in the midluteal phase and men ($N = 48$, unpublished data). In data set 6, the group manipulation was a laboratory stress induction (the MAST; Smeets et al., 2012) either immediately following acquisition or immediately prior to extinction ($N = 45$, Ney et al., 2018). In all data sets, participants had no neurological or cardiovascular illnesses, no history of head injury or loss of consciousness, no drug use, no heavy alcohol use, and no psychiatric illnesses, other than PTSD in data sets 1–3.

Given the goals and framework of this study, it is unlikely that variability in data collection methods (e.g., reinforcement ratio) or experimental manipulations would affect results. This is because the predictor variable in our study is the analysis method itself. As such, our primary concern was to produce data that reflected data obtained during real experiments wherein any effects observed were the differences between analysis strategies due to all data sets being tested by all strategies.

2.3 | Apparatus and data reduction

In all studies, a stimulus isolator (ADInstruments) was attached to the right hand and participants were encouraged to choose a US level that was "highly uncomfortable but not painful." The 500 ms electric shock was delivered at CS+ offset during the fear acquisition phase. Galvanic skin conductance was recorded in micro-Siemens (μS) using a 22 mVrms, 75 Hz constant-voltage coupler (ADInstruments). Electrodes were strapped to the second phalanges of the first and third fingers of the left hand. SCRs to the CS+ and CS− were preprocessed using the PsPM toolbox v4.2.1 in MATLAB (version 9.7) (Bach & Friston, 2013; Bach, Friston, & Dolan, 2013). Using custom coding, we used a peak scoring interval of 0.9–5 s following stimulus onset, given that SCRs peak within a relatively narrow window following CS onset, called the first interval response. However, this choice does not necessarily reflect a standardized latency interval as currently, this does not exist (see Jentsch, Wolf, & Merz, 2020; Pineles, Orr, & Orr, 2009). In order to remove noise in the data, a bidirectional Butterworth filter (1.5Hz low pass; 0.5Hz high pass) was applied to the raw SCR trace.

2.4 | Statistical analysis

In all analysis strategies, we aimed to test the stimulus × trial × group effects. For some methods, this meant that the analysis was actually a trial × group, or even phase × group interaction, since some methods used differential responses (calculated by subtracting a CS− response from the adjacent CS+ response) or averaged responses (either differential or CS+/CS− over successive trials, see Table 1). From each analysis we obtained a partial eta squared effect size for this interaction. Kendall non-parametric ranked order correlation coefficients ($_{T}b$) were run on the effect size from each data set for each of the three categories of analysis. Bayes factors and 95% credible intervals were calculated based on each correlation. This approach was favored over p values due to significant values being easily achieved in large sample sizes of simulated data. Further, credible intervals allow more accurate interpretation of the possible range of the effect size relative to confidence intervals (Morey, Hoekstra, Rouder, Lee, & Wagenmakers, 2016). To ensure that our resampled data sets did not bias the data, correlations were run and compared with both the original sample ($N = 6$, see Supporting Information) and the resampled sample ($N = 40$). All data analyses were conducted in Jamovi 1.1.9. Bayesian analyses were conducted using the jsq module.

3 | RESULTS

Over 5,000 unique articles were identified in the search. Fifteen articles were selected as they met the following

criteria: over 150 citations, fear conditioning and extinction phases, human only, and using skin conductance. Additional articles that had been cited less than 150 times were also included to increase the number of different methods examined (Strategies 3 and 4 in Acquisition-Extinction, Strategy 2 in Extinction, and Strategy 5 in Extinction-Extinction, Table 1). Therefore, this is a small, yet exemplary sample of the methods used in the fear conditioning literature.

As in Lonsdorf, Merz, et al. (2019), we observed a high heterogeneity of analytical strategies (Table 1). In Table 1, each strategy is assigned to a category based on how the phases were analyses (i.e., comparing acquisition-extinction, extinction as a whole or comparing early extinction-late extinction). The study that used each strategy is specified in the rightmost column. The differences between these strategies included how many trials were included in the study (column 3, Table 1), how many trials from these were included in the analysis (column 4, Table 1), whether these trials were averaged or assessed on a trial-by-trial basis (column 5, Table 1), whether the CS+/CS− trials were included as a single differential response (column 6, Table 1), and what final statistical method was used (column 7, Table 1). Different combinations of these variables lead to a potentially wide array of statistical strategies. We noted heterogeneity in the number of trials retained during the analysis, regardless of how many trials were originally present in the study. There was also inconsistency in whether selected trials were averaged or compared on a trial-by-trial basis, as well as whether differential responses were calculated. Resulting statistical analyses were more homogenous, with mixed ANOVAs being used across all high-impact studies.

3.1 | Acquisition-extinction

Strategies for the first set of analyses, where the change in responding from acquisition to extinction learning is assessed, were relatively similar (Table 1). All four strategies used average differential responses, and two of the four drew trials from the whole acquisition phase. One of the other strategies used the trials from the second half of acquisition, whereas the other strategy used the single highest differential response from acquisition. Two of the four strategies used the final two trials of extinction learning, one used the last three out of seven trials and the final used the first half of extinction trials.

3.2 | Static extinction

For the second set of analyses, we compared strategies from studies assessing extinction learning as a static construct (EXT) that could be compared to responses in other trials or studies. This group of strategies did not measure the change in responding across or within extinction learning phases and instead estimate the gross responding during extinction learning. Four out of seven compared CS+ and CS− responses, whereas the other three used differential responding. Three used trial-by-trial analyses; though of these, one used only the first two trials, one used all trials, and the final one used a "running average" response, where trials one and two were averaged as a single response, trials two and three were averaged, and so on. Three strategies used averaged responses, with one using the final quarter of extinction trials, one using the last half and one using the last two trials. Strategy 4 used only one trial; this was the last trial.

3.3 | Early extinction versus late extinction

For the final set of analyses, we compared the strategies from studies that assessed change in extinction learning across the extinction phase. Trial-by-trial analysis was not sufficient to fit this category, since ANOVA that fits trial as a parameter does not account for the order of the trials. Three of the five strategies compared the average of the first half of trials to the average of the second half of trials, though one of these strategies used differential responses, one only used CS+ responses, and the other retained the CS+ and CS− as separate responses. One of the strategies assessed, the average CS+ responses in the first quarter of extinction to the final quarter of extinction, and the final strategy assessed CS+ and CS− separate responses using linear trends across all trials.

3.4 | Correlations

Tables 2–4 show Kendall rank correlation coefficient values ($_{\tau}b$) for the three different sets of analyses. For strategies comparing acquisition and extinction phases, correlations were high between Strategies 1–3 (Table 2). Strategy 4 did not produce reliable results compared to other methods. For strategies producing a static estimate of extinction learning (Table 3), correlations were more inconsistent, ranging from $_{\tau}b = -.062$ to $_{\tau}b = .602$. Only seven comparisons between all combinations of the seven strategies produced correlations that were supported by Bayes factors and 95% credible intervals, though some of these were very highly supported. The final set of strategies performed similarly to acquisition, with six out of ten comparisons of the five strategies producing supported correlations. These correlations ranged from $_{\tau}b = .060$ to $_{\tau}b = .982$, with Strategy 1 and Strategy 2 being almost exactly similar, but Strategy 5 being dissimilar to all the other strategies.

TABLE 2 *Acquisition–extinction*. Strategy comparisons using the Kendall rank correlation coefficient between data sets with changes from acquisition to extinction learning phases estimated

		Strategy 2	Strategy 3	Strategy 4
Strategy 1	$_{\tau}b$	.609	.794	.125
	BF 95%CI	571814*** **[.76, .36]**	1.55E+10*** **[.90, .52]**	.4 [.32, −.09]
Strategy 2	$_{\tau}b$		.558	.044
	BF 95%CI		52991*** **[.71, .31]**	.2 [.24, −.16]
Strategy 3	$_{\tau}b$			.152
	BF 95%CI			.5 [.35, −.06]

$N = 40$ data sets with correlations comparing strategies conducted in all data sets. CIs that do not cross zero are bold.

Abbreviations: $_{\tau}b$, Spearman's R coefficient. 95%CIs, 95% credible intervals. BF = Bayes Factor

***BF > 30.

TABLE 3 *Static extinction*. Strategy comparisons using the Kendall rank correlation coefficient between data sets with a static extinction learning efficacy estimated

		Strategy 2	Strategy 3	Strategy 4	Strategy 5	Strategy 6	Strategy 7
Strategy 1	$_{\tau}b$	.047	.488	.252	.332	.075	.020
	BF 95%CI	.2 [.25, −.16]	2799*** **[.65, .25]**	3 **[.44, .03]**	17* **[.51, .10]**	.3 [.27, −.14]	.2 [.22, −.19]
Strategy 2	$_{\tau}b$		−.008	−.014	−.010	.602	.483
	BF 95%CI		.2 [.20, −.21]	.2 [.19, −.22]	.2 [.19, −.21]	408227*** **[.75, .35]**	2370*** **[.65, .24]**
Strategy 3	$_{\tau}b$			.102	.425	.012	.001
	BF 95%CI			.3 [.30, −.11]	283*** **[.60, .19]**	.2 [.21, −.19]	.2 [.20, −.20]
Strategy 4	$_{\tau}b$				.371	−.031	−.062
	BF 95%CI				51*** **[.55, .14]**	.2 [.17, −.23]	.2 [.14, −.26]
Strategy 5	$_{\tau}b$					−.052	.006
	BF 95%CI					.2 [.15, −.25]	.2 [.20, −.21]
Strategy 6	$_{\tau}b$						.152
	BF 95%CI						.5 [.34, −.06]

$N = 40$ data sets with correlations comparing strategies conducted in all data sets. CIs that do not cross zero are bold.

Abbreviations: BF, bayes factor, $_{\tau}b$, Spearman's R coefficient; 95%CIs, 95% credible intervals. BF = Bayes Factor

***BF > 30; *BF > 10.

TABLE 4 *Early–late extinction*. Strategy comparisons using the Kendall rank correlation coefficient between data sets with changes during extinction learning estimated

		Strategy 2	Strategy 3	Strategy 4	Strategy 5
Strategy 1	$_{\tau}b$	.982	.340	.295	.060
	BF 95%CI	4.89E+15*** **[.97, .67]**	21** **[.52, .11]**	7 **[.48, .07]**	.2 [.26, −.15]
Strategy 2	$_{\tau}b$		.358	.308	.068
	BF 95%CI		35*** **[.53, .13]**	9 **[.49, .08]**	.2 [.27, −.14]
Strategy 3	$_{\tau}b$			.630	.080
	BF 95%CI			1.53E+6*** **[.77, .38]**	.3 [.28, −.13]
Strategy 4	$_{\tau}b$				.083
	BF 95%CI				.3 [.28, −.13]

$N = 40$ data sets with correlations comparing strategies conducted in all data sets. CIs that do not cross zero are bold.

Abbreviations: $_{\tau}b$, Spearman's R coefficient. 95%CIs, 95% credible intervals. BF = Bayes Factor

***BF > 30; **BF > 20.

4 | DISCUSSION

Previous studies have reported high heterogeneity in the indexation and analysis of extinction retention and reinstatement between fear conditioning and extinction paradigms (Haaker et al., 2014; Lonsdorf, Merz, et al., 2019; Ney et al., 2018). In this study, we compared analytical strategies that assessed fear extinction learning in human SCR paradigms in several data sets that were resampled from our laboratory's data. A high degree of heterogeneity was found between the strategies, with choices such as which trials to use during analysis, whether to use differential responses and whether to average trials or use trial-by-trial analysis all differing significantly between studies. Using a bootstrapped data set based on six of our own data sets, we found that correlations between the strategies used in these studies were usually poor, even though they were intended to estimate similar constructs. We found this was true particularly for studies estimating SCRs both statically and across extinction learning, though strategies that assessed change between acquisition to extinction phases were relatively reliable. These findings have implications for the reliability of psychophysiological studies of fear acquisition and extinction learning.

When considering changes in SCRs from acquisition to extinction learning, strategies that compared average or maximal differential values during acquisition to average differential values at the end of extinction learning were highly correlated, regardless of the trials that were included. Strategy 4 of this category, which compared the average differential trials from late acquisition to early extinction was poorly correlated with the other strategies. We can surmise from this that it is likely that studies that compare different stages of each phase from acquisition to extinction may not be comparable. However, it should be noted that we had less acquisition trials compared to many other studies and this finding may not generalize to paradigms with more trials.

During extinction learning Strategies 1 and 3 were highly correlated, with the only difference being the inclusion of a quarter of the extinction trials. However, when strategies selected from different sections of extinction, they were poorly correlated. This was also reflected in the early late extinction category, with Strategies 3 and 4 being significantly correlated. This again suggests that analyses during extinction are relatively insensitive to minor variations in trial selection, so long as sufficiently large numbers of trials are selected from the same quadrants of the phase. Using linear trends rather than omnibus ANOVA resulted in vastly different effect sizes. Interestingly, the evidence here also shows that the use of differential compared to separate CS+/CS− responding may not impact robustness, with high correlations observed in both Categories 2 and 3 between studies that used identical parameters apart from this. It can, therefore, be concluded, based on these data and with relatively homogenous trial numbers between studies, that selection of trials from contrasting segments of paradigm phases and discrepant use of trial-by-trial compared to averaged data present the major risks to robustness.

We have previously made several recommendations that may improve robustness in the fear conditioning paradigm (Ney et al., 2018). Here, we maintain that graphing trial-by-trial data and increasing sample size are ways to improve transparency and robustness that any laboratory should be readily able to implement with minimal effort and resources. Similarly, the transparency and robustness of research might be improved by any laboratory by adopting a multiverse approach, where multiple analyses are conducted on the same data to elicit the reliability of reported findings from one approach (Silberzahn et al., 2018; Steegen, Tuerlinckx, Gelman, & Vanpaemel, 2016). These approaches rely on increased transparency in data reporting and analysis, and we maintain that decisions during data reduction and analysis should be reported and justified (Lonsdorf, Klingelhöfer-Jens, et al., 2019; Ney et al., 2018). It is also possible that reproducibility may be improved by computational optimization of paradigm design, which may aid in determining how to vary experimental parameters to answer specific research questions (Melinscak & Bach, 2020).

Based on the current data, however, we make several specific recommendations that may improve robustness. First, future research should recognize that learning between early and late stages of an extinction phase is unlikely to be comparable, since the differential selection of these time periods presented the greatest impairment to robustness in the present study. Future studies should aim to specify and further characterize the differences in learning that occur in early compared to late extinction trials so that it can be better understood at which stage of learning clinical participants show the most impairment or treatments show the most efficacy. Similarly, the cause for inadequate robustness between trial-by-trial and averaged data should be systematically investigated. It is possible that the failure of these methods to replicate is due to lack of power, in which case methods that seek to improve power via experimental design and SCR scoring are highly desirable (Bach & Melinscak, 2020; Melinscak & Bach, 2020). Specifically, improving SCR scoring can provide better estimates of responses relevant to extinction paradigms by reducing measurement error (Bach & Melinscak, 2020), and optimization of experimental designs based on statistical requirements can make analyses more amenable to experimental data (Melinscak & Bach, 2020). Both of these methods can improve experimental power without additional participant recruitment.

While reducing measurement error during pre-processing of data can improve power, a greater understanding of the mechanisms that shape fear extinction learning could

also be achieved through the implementation of computational learning models of processed data. Model-based analysis has previously been used to characterize dissociable striatal and amygdala contributions to fear conditioning (Delgado, Li, Schiller, & Phelps, 2008; Li, Schiller, Schoenbaum, Phelps, & Daw, 2011; Schiller, Levy, Niv, LeDoux, & Phelps, 2008), accounting for genetic, affective, and cognitive individual differences in fear learning (Baetu et al., 2018; Laing, Burns, & Baetu, 2019), and identifying exaggerated neural prediction errors in PTSD symptomology (Homan et al., 2019). Tzovara, Korn, and Bach (2018) recently found that both SCRs and pupil responses during conditioning were best explained by a Bayesian learning model, though reflected slightly different aspects of learning during the task. However, these models, as well as Bayesian learning models that parameterize uncertainty (Gershman & Hartley, 2015; Tzovara et al., 2018), have thus far only been applied to human fear conditioning in a limited way. Computational modeling is advantageous because it formalizes statistical decision-making to allow statistical choices between studies to be explicitly compared. Comparison and selection of best statistical methods require complete transparency of justifications for different choices and aims to formally determine the most representative model for a particular experimental design. For this reason, modeling adheres very closely to the goals of open science and represents the best practices in statistical analysis (Adams, Huys, & Roiser, 2016). Therefore, by mathematically expressing what we expect to occur in SCR data during extinction learning, we can assess the validity of our existing ideas of how extinction learning impacts physiological response. With this level of formalism, methodological decisions are required to be justified and compared to competing frameworks.

One limitation of the current study is that the level of heterogeneity found here may not generalize to other data processing methods, such as model-fitting techniques such as PsPM where study power is maximized (Bach & Melinscak, 2020). Further, significant work will need to be conducted before standardization of statistical analyses of this paradigm may be achieved; here we have only indicated that systemic issues exist in the current approach. Modeling approaches will also need to be tailored to suit different paradigm designs to accommodate parameters such as trial length (Bach & Melinscak, 2020). Our paradigm also featured a relatively small number of trials, particularly during acquisition. Therefore, our conclusions of high correlations between certain methods may not generalize to studies with more trials and cannot be generalized to different experimental designs such as delayed extinction. Finally, due to the high heterogeneity of strategies anticipated in a literature search, our included studies were compiled to provide an exemplary, yet non-exhaustive, representation of strategies used in the field. Although many of the studies here were chosen on the basis of the number of citations rather than statistical methods, larger syntheses of the literature may reveal less heterogeneity and higher robustness than we observed here. Similarly, differences in the original data included here as derived from different studies may have caused higher variance in the data and potentially impacted our results.

In summary, we provide evidence of limited robustness between SCR fear extinction studies due to variation in analytical strategy. The highest impact on robustness was evidenced by differential trial selection from contrasting halves of extinction learning, as well as the use of trial-by-trial compared to averaged analyses. We conclude that, in order to enhance reliability, future studies should investigate the differences in extinction learning that occurs between early and late extinction phases. We have also made several suggestions that could improve the robustness of the current paradigm, including improving sample sizes, visualizing SCR data points, improving transparency of data reporting, and computational modeling of extinction learning.

ACKNOWLEDGMENTS

This work was supported by an NHMRC Program grant to KLF (APP1073041).

CONFLICT OF INTEREST

The authors have no conflicts of interest to report.

ORCID

Luke John Ney https://orcid.org/0000-0003-0209-8366

REFERENCES

Adams, R. A., Huys, Q. J. M., & Roiser, J. P. (2016). Computational psychiatry: Towards a mathematically informed understanding of mental illness. *Journal of Neurology, Neurosurgery & Psychiatry*, *87*(1), 53. https://doi.org/10.1136/jnnp-2015-310737

Bach, D., Castegnetti, G., Korn, C. W., Gerster, S., Melinscak, F., & Moser, T. (2018). Psychophysiological modeling: Current state and future directions. *Psychophysiology*, *55*(11), e13214. https://doi.org/10.1111/psyp.13209

Bach, D., & Friston, K. J. (2013). Model-based analysis of skin conductance responses: Towards causal models in psychophysiology. *Psychophysiology*, *50*(1), 15–22. https://doi.org/10.1111/j.1469-8986.2012.01483.x

Bach, D., Friston, K. J., & Dolan, R. J. (2013). An improved algorithm for model-based analysis of evoked skin conductance responses. *Biological Psychology*, *94*(3), 490–497. https://doi.org/10.1016/j.biopsycho.2013.09.010

Bach, D., & Melinscak, F. (2020). Psychophysiological modelling and the measurement of fear conditioning. *Behaviour Research and Therapy*, *127*, 103576. https://doi.org/10.1016/j.brat.2020.103576

Baetu, I., Pitcher, J. B., Cohen-Woods, S., Lancer, B., Beu, N., Foreman, L. M., … Burns, N. R. (2018). Polymorphisms that affect GABA neurotransmission predict processing of aversive prediction errors

in humans. *NeuroImage*, *176*, 179–192. https://doi.org/10.1016/j.neuroimage.2018.04.058

Blechert, J., Michael, T., Vriends, N., Margraf, J., & Wilhelm, F. H. (2007). Fear conditioning in posttraumatic stress disorder: Evidence for delayed extinction of autonomic, experiential, and behavioural responses. *Behavior Research and Therapy*, *45*(9), 2019–2033. https://doi.org/10.1016/j.brat.2007.02.012

Boucsein, W. (2012). *Electrodermal activity* (2nd ed.) New York, NY: Springer.

Bouton, M. E. (2004). Context and behavioral processes in extinction. *Learn Mem*, *11*(5), 485–494. https://doi.org/10.1101/lm.78804

Craske, M. G., Rauch, S. L., Ursano, R., Prenoveau, J., Pine, D. S., & Zinbarg, R. E. (2009). What is an anxiety disorder? *Depress Anxiety*, *26*(12), 1066–1085. https://doi.org/10.1002/da.20633

Craske, M. G., Treanor, M., Conway, C. C., Zbozinek, T., & Vervliet, B. (2014). Maximizing exposure therapy: An inhibitory learning approach. *Behaviour Research and Therapy*, *58*, 10–23. https://doi.org/10.1016/j.brat.2014.04.006

Delgado, M. R., Li, J., Schiller, D., & Phelps, E. A. (2008). The role of the striatum in aversive learning and aversive prediction errors. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *363*(1511), 3787–3800. https://doi.org/10.1098/rstb.2008.0161

Felmingam, K. L., Ney, L. J., Caruana, J. M., Miller, L. N., Zuj, D. V., Hsu, C. M., Bryant, R. (under review). Lower estradiol predicts increased reinstatement of fear in women.

Garfinkel, S. N., Abelson, J. L., King, A. P., Sripada, R. K., Wang, X., Gaines, L. M., & Liberzon, I. (2014). Impaired contextual modulation of memories in PTSD: An fMRI and psychophysiological study of extinction retention and fear renewal. *Journal of Neuroscience*, *34*(40), 13435–13443. https://doi.org/10.1523/JNEUROSCI.4287-13.2014

Gershman, S. J., & Hartley, C. A. (2015). Individual differences in learning predict the return of fear. *Learn Behav*, *43*(3), 243–250. https://doi.org/10.3758/s13420-015-0176-z

Grady, A. K., Bowen, K. H., Hyde, A. T., Totsch, S. K., & Knight, D. C. (2016). Effect of continuous and partial reinforcement on the acquisition and extinction of human conditioned fear. *Behavioral Neuroscience*, *130*(1), 36–43. https://doi.org/10.1037/bne0000121

Graham, B. M., Callaghan, B. L., & Richardson, R. (2014). Bridging the gap: Lessons we have learnt from the merging of psychology and psychiatry for the optimisation of treatments for emotional disorders. *Behavior Research and Therapy*, *62*, 3–16. https://doi.org/10.1016/j.brat.2014.07.012

Graham, B. M., & Milad, M. R. (2013). Blockade of estrogen by hormonal contraceptives impairs fear extinction in female rats and women. *Biological Psychiatry*, *73*(4), 371–378. https://doi.org/10.1016/j.biopsych.2012.09.018

Grupe, D. W., & Nitschke, J. B. (2013). Uncertainty and anticipation in anxiety: An integrated neurobiological and psychological perspective. *Nature Reviews Neuroscience*, *14*(7), 488–501. https://doi.org/10.1038/nrn3524

Haaker, J., Golkar, A., Hermans, D., & Lonsdorf, T. B. (2014). A review on human reinstatement studies: an overview and methodological challenges. *Learning & Memory*, *21*(9), 424–440. https://doi.org/10.1101/lm.036053.114

Homan, P., Levy, I., Feltham, E., Gordon, C., Hu, J., Li, J., … Schiller, D. (2019). Neural computations of threat in the aftermath of combat trauma. *Nature Neuroscience*, *22*(3), 470–476. https://doi.org/10.1038/s41593-018-0315-x

Ioannidis, J. P. (2005). Why most published research findings are false. *PLoS Med*, *2*(8), e124. https://doi.org/10.1371/journal.pmed.0020124

Jentsch, V. L., Wolf, O. T., & Merz, C. J. (2020). Temporal dynamics of conditioned skin conductance and pupillary responses during fear acquisition and extinction. *International Journal of Psychophysiology*, *147*, 93–99. https://doi.org/10.1016/j.ijpsycho.2019.11.006

Laing, P. A. F., Burns, N., & Baetu, I. (2019). Individual differences in anxiety and fear learning: The role of working memory capacity. *Acta Psychologica*, *193*, 42–54. https://doi.org/10.1016/j.actpsy.2018.12.006

Lebois, L. A. M., Seligowski, A. V., Wolff, J. D., Hill, S. B., & Ressler, K. J. (2019). Augmentation of extinction and inhibitory learning in anxiety and trauma-related disorders. *Annual Review of Clinical Psychology*, *15*, 257–284. https://doi.org/10.1146/annurev-clinpsy-050718-095634

Li, J., Schiller, D., Schoenbaum, G., Phelps, E. A., & Daw, N. D. (2011). Differential roles of human striatum and amygdala in associative learning. *Nature Neuroscience*, *14*(10), 1250–1252. https://doi.org/10.1038/nn.2904

Lonsdorf, T. B., Klingelhöfer-Jens, M., Andreatta, M., Beckers, T., Chalkia, A., Gerlicher, A., … Merz, C. J. (2019). Navigating the garden of forking paths for data exclusions in fear conditioning research. *Elife*, *8*, e52465. https://doi.org/10.7554/eLife.52465

Lonsdorf, T. B., Menz, M. M., Andreatta, M., Fullana, M. A., Golkar, A., Haaker, J., … Merz, C. J. (2017). Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear. *Neuroscience and Biobehavioral Reviews*, *77*, 247–285. https://doi.org/10.1016/j.neubiorev.2017.02.026

Lonsdorf, T. B., Merz, C. J., & Fullana, M. A. (2019). Fear extinction retention: Is it what we think it is? *Biological Psychiatry*, *85*(12), 1074–1082. https://doi.org/10.1016/j.biopsych.2019.02.011

Lovibond, P. F., Mitchell, C. J., Minard, E., Brady, A., & Menzies, R. G. (2009). Safety behaviours preserve threat beliefs: Protection from extinction of human fear conditioning by an avoidance response. *Behaviour Research and Therapy*, *47*(8), 716–720. https://doi.org/10.1016/j.brat.2009.04.013

Melinscak, F., & Bach, D. R. (2020). Computational optimization of associative learning experiments. *PLOS Computational Biology*, *16*(1), e1007593. https://doi.org/10.1371/journal.pcbi.1007593

Michael, T., Blechert, J., Vriends, N., Margraf, J., & Wilhelm, F. H. (2007). Fear conditioning in panic disorder: Enhanced resistance to extinction. *Journal of Abnormal Psychology*, *116*(3), 612–617. https://doi.org/10.1037/0021-843x.116.3.612

Milad, M. R., Furtak, S. C., Greenberg, J. L., Keshaviah, A., Im, J. J., Falkenstein, M. J., … Wilhelm, S. (2013). Deficits in conditioned fear extinction in obsessive-compulsive disorder and neurobiological changes in the fear circuit. *JAMA Psychiatry*, *70*(6), 608–618. https://doi.org/10.1001/jamapsychiatry.2013.914

Milad, M. R., Goldstein, J. M., Orr, S. P., Wedig, M. M., Klibanski, A., Pitman, R. K., & Rauch, S. L. (2006). Fear conditioning and extinction: Influence of sex and menstrual cycle in healthy humans. *Behavioral Neuroscience*, *120*(6), 1196–1203. https://doi.org/10.1037/0735-7044.120.5.1196

Milad, M. R., Orr, S. P., Lasko, N. B., Chang, Y., Rauch, S. L., & Pitman, R. K. (2008). Presence and acquired origin of reduced recall for fear extinction in PTSD: Results of a twin study. *Journal of Psychiatric Research*, *42*(7), 515–520. https://doi.org/10.1016/j.jpsychires.2008.01.017

Milad, M. R., Pitman, R. K., Ellis, C. B., Gold, A. L., Shin, L. M., Lasko, N. B., … Rauch, S. L. (2009). Neurobiological basis of failure to recall extinction memory in posttraumatic stress disorder. *Biological Psychiatry*, *66*(12), 1075–1082. https://doi.org/10.1016/j.biopsych.2009.06.026

Milad, M. R., & Quirk, G. J. (2012). Fear extinction as a model for translational neuroscience: Ten years of progress. *Annual Review of Psychology*, *63*, 129–151. https://doi.org/10.1146/annurev.psych.121208.131631

Milad, M. R., Zeidan, M. A., Contero, A., Pitman, R. K., Klibanski, A., Rauch, S. L., & Goldstein, J. M. (2010). The influence of gonadal hormones on conditioned fear extinction in healthy humans. *Neuroscience*, *168*(3), 652–658. https://doi.org/10.1016/j.neuroscience.2010.04.030

Morey, R. D., Hoekstra, R., Rouder, J. N., Lee, M. D., & Wagenmakers, E.-J. (2016). The fallacy of placing confidence in confidence intervals. *Psychonomic Bulletin & Review*, *23*(1), 103–123. https://doi.org/10.3758/s13423-015-0947-8

Ney, L. J., Nicholson, E., Nichols, D., Felmingam, K. L., Bruno, R., & Matthews, A. (in prep). Endocannabinoids during fear conditioning, extinction and extinction recall in PTSD.

Ney, L. J., Wade, M., Reynolds, A., Zuj, D. V., Dymond, S., Matthews, A., & Felmingham, K. L. (2018). Critical evaluation of current data analysis strategies for psychophysiological measures of fear conditioning and extinction in humans. *International Journal of Psychophysiology*, *134*, 95–107. https://doi.org/10.1016/j.ijpsycho.2018.10.010

Ojala, K. E., & Bach, D. R. (2020). Measuring learning in human classical threat conditioning: Translational, cognitive and methodological considerations. *Neuroscience and Biobehavioral Reviews*, *114*, 96–112. https://doi.org/10.1016/j.neubiorev.2020.04.019

Pace-Schott, E. F., Spencer, R. M. C., Vijayakumar, S., Ahmed, N. A. K., Verga, P. W., Orr, S. P., … Milad, M. R. (2013). Extinction of conditioned fear is better learned and recalled in the morning than in the evening. *Journal of Psychiatric Research*, *47*(11), 1776–1784. https://doi.org/10.1016/j.jpsychires.2013.07.027

Phelps, E. A., Delgado, M. R., Nearing, K. I., & LeDoux, J. E. (2004). Extinction learning in humans: Role of the amygdala and vmPFC. *Neuron*, *43*(6), 897–905. https://doi.org/10.1016/j.neuron.2004.08.042

Pineles, S. L., Orr, M. R., & Orr, S. P. (2009). An alternative scoring method for skin conductance responding in a differential fear conditioning paradigm with a long-duration conditioned stimulus. *Psychophysiology*, *46*(5), 984–995. https://doi.org/10.1111/j.1469-8986.2009.00852.x

Pittig, A., Treanor, M., LeBeau, R. T., & Craske, M. G. (2018). The role of associative fear and avoidance learning in anxiety disorders: Gaps and directions for future research. *Neuroscience and Biobehavioral Reviews*, *88*, 117–140. https://doi.org/10.1016/j.neubiorev.2018.03.015

Schiller, D., Levy, I., Niv, Y., LeDoux, J. E., & Phelps, E. A. (2008). From fear to safety and back: Reversal of fear in the human brain. *Journal of Neuroscience*, *28*(45), 11517–11525. https://doi.org/10.1523/jneurosci.2265-08.2008

Schiller, D., Monfils, M. H., Raio, C. M., Johnson, D. C., Ledoux, J. E., & Phelps, E. A. (2010). Preventing the return of fear in humans using reconsolidation update mechanisms. *Nature*, *463*(7277), 49–53. https://doi.org/10.1038/nature08637

Silberzahn, R., Uhlmann, E. L., Martin, D. P., Anselmi, P., Aust, F., Awtrey, E., … Nosek, B. A. (2018). Many analysts, one data set: Making transparent how variations in analytic choices affect results. *Advances in Methods and Practices in Psychological Science*, *1*(3), 337–356. https://doi.org/10.1177/2515245917747646

Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2011). False-positive psychology: Undisclosed flexibility in data collection and analysis allows presenting anything as significant. *Psychological Science*, *22*(11), 1359–1366. https://doi.org/10.1177/0956797611417632

Simon, J., Bruce, P., & Troiana, V. (2013). *Resampling stats add-in for excel*. Arlington, Virginia: Statistics.com.

Smeets, T., Cornelisse, S., Quaedflieg, C. W., Meyer, T., Jelicic, M., & Merckelbach, H. (2012). Introducing the Maastricht Acute Stress Test (MAST): A quick and non-invasive approach to elicit robust autonomic and glucocorticoid stress responses. *Psychoneuroendocrinology*, *37*(12), 1998–2008. https://doi.org/10.1016/j.psyneuen.2012.04.012

Soliman, F., Glatt, C. E., Bath, K. G., Levita, L., Jones, R. M., Pattwell, S. S., … Casey, B. J. (2010). A genetic variant BDNF polymorphism alters extinction learning in both mouse and human. *Science*, *327*(5967), 863–866. https://doi.org/10.1126/science.1181886

Steegen, S., Tuerlinckx, F., Gelman, A., & Vanpaemel, W. (2016). Increasing transparency through a multiverse analysis. *Perspectives on Psychological Science*, *11*(5), 702–712. https://doi.org/10.1177/1745691616658637

Suarez-Jimenez, B., Albajes-Eizagirre, A., Lazarov, A., Zhu, X. I., Harrison, B. J., Radua, J., … Fullana, M. A. (2019). Neural signatures of conditioning, extinction learning, and extinction recall in posttraumatic stress disorder: A meta-analysis of functional magnetic resonance imaging studies. *Psychological Medicine*, 1–10. https://doi.org/10.1017/S0033291719001387

Tzovara, A., Korn, C. W., & Bach, D. (2018). Human Pavlovian fear conditioning conforms to probabilistic learning. *PLOS Computational Biology*, *14*(8), e1006243. https://doi.org/10.1371/journal.pcbi.1006243

Vicario, C. M., Nitsche, M. A., Hoysted, I., Yavari, F., Avenanti, A., Salehinejad, M. A., & Felmingham, K. L. (2020). Anodal transcranial direct current stimulation over the ventromedial prefrontal cortex enhances fear extinction in healthy humans: A single blind sham-controlled study. *Brain Stimulation*, *13*(2), 489–491. https://doi.org/10.1016/j.brs.2019.12.022

White, E. C., & Graham, B. M. (2016). Estradiol levels in women predict skin conductance response but not valence and expectancy ratings in conditioned fear extinction. *Neurobiology of Learning and Memory*, *134*, 339–348, https://doi.org/10.1016/j.nlm.2016.08.011

Zeidan, M. A., Igoe, S. A., Linnman, C., Vitalo, A., Levine, J. B., Klibanski, A., … Milad, M. R. (2011). Estradiol modulates medial prefrontal cortex and amygdala activity during fear extinction in women and female rats. *Biological Psychiatry*, *70*(10), 920–927. https://doi.org/10.1016/j.biopsych.2011.05.016

Zuj, D. V., & Norrholm, S. D. (2019). The clinical applications and practical relevance of human conditioning paradigms for posttraumatic stress disorder. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *88*, 339–351. https://doi.org/10.1016/j.pnpbp.2018.08.014

Zuj, D. V., Palmer, M. A., Hsu, C. M., Nicholson, E. L., Cushing, P. J., Gray, K. E., & Felmingham, K. L. (2016). Impaired fear extinction associated with PTSD increases with hours-since-waking. *Depress Anxiety*, *33*(3), 203–210. https://doi.org/10.1002/da.22463

Zuj, D. V., Palmer, M. A., Lommen, M. J., & Felmingham, K. L. (2016). The centrality of fear extinction in linking risk factors to PTSD: A narrative review. *Neuroscience and Biobehavioral Reviews*, *69*, 15–35. https://doi.org/10.1016/j.neubiorev.2016.07.014

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.
Supplementary Material

TABLE S1 Summary of differences in datasets
TABLE S2 *Acquisition–extinction*. Strategy comparisons using Kendall rank correlation coefficient between datasets with changes from acquisition to extinction learning phases estimated
TABLE S3 *Static extinction*. Strategy comparisons using Kendall rank correlation coefficient between datasets with a static extinction learning efficacy estimated
TABLE S4 *Early–late extinction*. Strategy comparisons using Kendall rank correlation coefficient between datasets with changes during extinction learning estimated

How to cite this article: Ney LJ, Laing PAF, Steward T, Zuj DV, Dymond S, Felmingham KL. Inconsistent analytic strategies reduce robustness in fear extinction via skin conductance response. *Psychophysiology*. 2020;00:e13650. https://doi.org/10.1111/psyp.13650