Learning & Behavior

Extinction in multiple contexts reduces the return of extinguished responses: A multilevel meta-analysis

Journal:	*Learning & Behavior*
Manuscript ID	LB-OR-23-062
Manuscript Type:	Original
Date Submitted by the Author:	29-Jun-2023
Complete List of Authors:	Bustamante, Javier; Universidad de O'Higgins Soto, Marcela; Pontificia Universidad Catolica de Chile Miguez, Gonzalo; Universidad de Chile Quezada Scholz, Vanetza; Universidad de Chile Angulo, Rocío; Universidad de O'Higgins Laborda, Mario; Universidad de Chile
Keywords:	Exposure therapy, extinction, Extinction in multiple contexts, Recovery, Relapse

Extinction in multiple contexts reduces the return of extinguished responses: A multilevel meta-analysis

Javier Bustamante[a], Marcela Soto[b], Gonzalo Miguez[c], Vanetza E. Quezada-Scholz[c], Rocío Angulo[a], & Mario A. Laborda[c,*]

[a]Institute of Social Sciences, Universidad de O'Higgins, Av. Libertador Bernardo O'Higgins 611, Rancagua, Chile. javier.bustamante@uoh.cl, rocio.angulo@uoh.cl

[b]School of Psychology, Pontificia Universidad Católica de Chile, Av. Vicuña Mackenna 4860, Santiago, Chile. misoto5@uc.cl

[c]Department of Psychology, Universidad de Chile, Av. Ignacio Carrera Pinto 1045, Santiago, Chile. gonzalo_miguez@uchile.cl, qvanetza@u.uchile.cl, mariolaborda@u.uchile.cl

Corresponding author: Mario A. Laborda
Mailing Address: Department of Psychology
Universidad de Chile
Av. Ignacio Carrera Pinto 1045
Ñuñoa, Santiago, Chile.
Postal Code 7800284
Telephone number: +56 9 92777097
E-mail: mariolaborda@u.uchile.cl

Abstract

Extinguished responses have been shown to reappear under several circumstances, which is considered to model relapse after exposure therapy. Conducting extinction in multiple contexts has been explored as a technique to decrease the recovery of extinguished responses. The present meta-analysis aimed to examine whether extinction in multiple contexts can consistently reduce the recovery of extinguished responses. After searching in several databases, experiments were included in the analysis if they presented extinction in multiple contexts, an experimental design, and an adequate statistical report. Cohen's *d* was obtained for each critical comparison and weighted to obtain the sample's average weighted effect size. Analyses were then performed using a multilevel meta-analytic approach. Twenty-five studies were included, with a total sample of 37 experiments or critical comparisons. The analyses showed a large effect size for the sample, moderated by the length of CS exposure, type of experimental subject, and type of recovery. The robust effect of extinction in multiple contexts on relapse should encourage clinicians to consider extinction in multiple contexts as a useful technique in therapy and research.

Keywords: Extinction, Exposure Therapy, Extinction in Multiple Contexts, Recovery, Relapse.

Extinction in multiple contexts reduces the return of extinguished responses: A multilevel meta-analysis

Extinction learning was first described by Ivan P. Pavlov at the early stages of the experimental study of classical conditioning (Pavlov, 1927). In his extinction experiments, Pavlov found that after conditioned salivation (i.e., conditioned response, CR) to an auditory cue has been established by pairing the auditory stimulus (i.e., the conditioned stimulus, CS) with food (i.e., unconditioned stimulus, US), this conditioned responding decreased when the auditory cue was repeatedly presented in the absence of food (i.e., extinction took place). Since then, the generality of extinction and its related phenomena have been widely studied for theoretical (e.g., Bouton, 1993; Bouton et al., 2021; Delamater, 2004; Delamater & Westbrook, 2014; McConnell et al., 2013; McConnell & Miller, 2012; Miguez, Witnauer, et al., 2014; Miller & Laborda, 2011; Nihei et al., 2020; Uengoer et al., 2020) and translational as well as applied reasons (e.g., Bowers & Ressler, 2015; Crombie et al., 2021; Lipp et al., 2020; O'Malley & Waters, 2018; Quezada et al., 2018; San Martín et al., 2018; Scheveneels et al., 2019; Spix et al., 2021; Waters et al., 2021; Zbozinek & Craske, 2018).

Regarding the potential applications of extinction, in the last couple of decades, a good deal of attention has been given to the study of extinction as a model for exposure therapy, since extinction learning is likely the most relevant basic phenomenon involved in this type of treatment (e.g., Craske et al., 2018; Waters et al., 2021). Just as with extinction in the laboratory, conditioned responding decreases in exposure therapy when patients are repeatedly confronted with situations and stimuli that, for example, have been previously associated with aversive consequences or with the effect of drugs in the organism (e.g., Laborda & Miguez, 2015; Mellentin et al., 2017). In the case of anxiety-related disorders (Craske et al., 2018), fear and anxiety responses decrease as patients are systematically confronted with their feared stimuli and situations, in the absence of their expected aversive consequences (e.g., exposure to the image of a dog for a phobic patient in the absence of its feared biting). Exposure therapy is the most effective therapeutic approach to most anxiety disorders (e.g., Chambless & Ollendick, 2001; Parker et al., 2018). However, in certain situations, we can expect relapse after treatment, a situation that in the case of fear and anxiety has been denominated return of fear (e.g., Rachman, 1989).

In the experimental study of extinction, several situations that resemble and model the return of extinguished responses after exposure therapy have been described (e.g., Bouton, 2002; Vervliet, Craske, et al., 2013; Vurbic & Bouton, 2014). In a spontaneous recovery situation, extinguished responses return when they are tested a time after extinction treatment (e.g., Pavlov, 1927; Rescorla, 2004), and in a renewal situation, extinguished responses return when they are tested out of the context in which extinction took place (e.g., Alfaro et al., 2018; Bouton & Bolles, 1979; Bouton & King, 1983; Polack et al., 2011). For example, if the acquisition of an association occurs in one context (A), extinction in a second context (B), and testing takes place back in the acquisition context (A), we are referring to an ABA renewal design. If instead testing occurs in a third context (context C), we are referring to an ABC renewal design.

Many behavioral manipulations have been developed to prevent response recovery after experimental extinction, with the expectation that they could also hold translational value and are potentially able to reduce relapse after exposure therapy (e.g., Bouton, Woods, et al., 2006; Laborda et al., 2011; Lipp et al., 2020; also see Laborda et al., 2014 for similar manipulations applied to fear prevention). However, not all reports have been positive concerning the effectiveness of these techniques (e.g., Bouton, García-Gutiérrez, et al., 2006; Bustamante et al., 2019; Neumann, 2006).

Based on the human learning literature, in which training an association in multiple contexts has shown to improve its recovery (e.g., Smith, 1982, 1985; Smith et al., 1978; for an example in conditioning without extinction see Miguez, Laborda, et al., 2014), Bouton (1991) suggested that training extinction in several different contexts, may encourage the potential generalization of extinction learning to other circumstances or contexts, and thus help decrease post-extinction recovery and relapse. Gunther et al. (1998) were the first to empirically evaluate whether extinction in multiple contexts does reduce recovery after extinction training, in a study conducted with rats. In their first experiment, subjects in the experimental group received fear conditioning in one context (A), extinction in three different contexts (BCD), and were finally tested for fear responding in yet another new context (E), that is, an ABC renewal design with extinction in multiple contexts. This group showed significantly less fear recovery than subjects in the control group, which received the same treatment, but with extinction in only one context (B), demonstrating for the first time that extinction in multiple contexts can reduce the recovery of extinguished responding, in this case preventing the ABC renewal of extinguished conditioned fear. In humans, this technique was first reported by Pineño and Miller (2004), using a predictive learning task. In their first experiment, participants had to avoid “mines” that were predicted by different lights as they were driving a truck with war refugees. Different town names were used as different contexts. For one group, the predictive relationship was trained in one context (A), extinguished in a different one (B), and then tested in a third context (C), while a second group received similar acquisition in one context (A), extinction in three different ones (BCD), and finally testing in a different context (E). Finally, a third group received similar training in only one context (A). The results showed that there was ABC renewal in humans, as well as attenuation of recovery after the extinction of predictive learning in multiple contexts, complementing the work that had been developed in non-human animals.

Since then, numerous studies have evaluated whether extinction in multiple contexts could prevent the recovery of extinguished conditioned responses, in different preparations and across different research groups and laboratories (e.g., Bernal-Gamboa et al., 2020; Chelonis et al., 1999; Glautier et al., 2013; González et al., 2016; Olatunji et al., 2017; Shiban et al., 2013). However, some reports have also failed to observe an effect of the treatment on response recovery (e.g., Bouton et al., 2006; Hermann et al., 2020; Neumann, 2006). Thus, it is of interest to integrate and review the existing evidence in order to examine whether extinction in multiple contexts is an effective tool for preventing relapse, and which elements present in the different studies might affect this effectiveness.

Even when some researchers have presented partial narrative syntheses concerning this and other techniques to reduce recovery after extinction (e.g., Laborda et al., 2011; Pittig et al., 2016; Weisman & Rodebaugh, 2018), to date, no meta-analysis nor systematic review on this effect has been conducted. Recently, Chao and colleagues (2021), have posted a preprint of a protocol aimed to systematically review and integrate the effect of extinction in multiple contexts, but focusing solely on fear recovery; the results to our knowledge have not been published yet.

Therefore, the aim of the present study was to review and integrate, using a multilevel meta-analytic approach, the evidence regarding the effect of conducting extinction in multiple contexts across all available experimental preparations, settings, and experimental subjects and participants. The analysis was conducted by estimating and integrating the effect sizes of the studies and calculating confidence intervals, using a multilevel meta-analytic approach (Van den Noortgate et al., 2015; Fernández-Castilla et al., 2020). We hypothesized, based on the available evidence, that extinction in multiple contexts would be effective at reducing response recovery or relapse, although potentially sensitive to differences in the experimental preparations used to examine it.

The protocol for the present study was not published before the review was conducted, but it is now available as Supplementary Material.

Method

The present research synthesis was performed following PRISMA and APA recommendations for conducting and reporting meta-analytical studies (Appelbaum et al., 2018; Cooper, 2017, 2018; Moher et al., 2009; Page et al., 2021).

Search Strategy

The following electronic databases were searched: Web of Knowledge (formerly Web of Science), Scopus, and APA PsycInfo. Additionally, two further searches were implemented in search for grey literature at PsyArXiv Preprints and the ProQuest Dissertations and Theses Global Database. Finally, the reference lists and previous citations of all selected articles were searched, and all corresponding authors were contacted via email and asked for both unpublished and published results. The final search was performed in June 2023.

Search terms were derived from existing literature on extinction in multiple contexts, which was previously known by the authors. The obtained search terms were compared and extended using the online APA Thesaurus of Psychological Index Terms (available on https://www.apa.org/pubs/databases/training/thesaurus), in order to assess their accuracy and obtain potential synonyms. The critical search terms were: ‘extinction’ AND ‘multiple contexts’

AND ‘recovery’, but we also included variants on each of these terms (Extinction: ‘exposure’ OR ‘confrontation’ OR ‘outcome interference’; multiple contexts: ‘several contexts’ OR ‘many contexts’; recovery: ‘ABC’ OR ‘ABA’ OR ‘AAB’ OR ‘AAC’ OR ‘renewal’ OR ‘spontaneous recovery’ OR ‘reinstatement’ OR ‘rapid reacquisition’ OR ‘relapse’ OR ‘return’ OR ‘resurgence’ OR ‘reoccurrence’ OR ‘context shift’ OR ‘context change’). Truncated search terms were used when corresponding (i.e., ‘extin*’ instead of ‘extinction’). For the specific search terms used in each database see the Supplementary Material.

Study Selection and Eligibility Criteria

Two independent researchers (JB and MS) selected the studies for the analysis, screening titles and abstracts against the eligibility criteria. In the cases that were deemed relevant or ambiguous, the whole text was screened. After the first independent screening, both reviewers met to compare and resolve discrepancies. In all cases, the final eligibility decision was made by four researchers (JB, MS, GM, and ML).

Studies were selected for the present meta-analysis if they a) presented extinction training in multiple contexts; b) were written in either English or Spanish; c) presented a post-extinction recovery assessment (e.g., renewal, spontaneous recovery); d) used an experimental design (i.e., presence of experimental and control conditions, manipulation of an independent variable, and random assignment of subjects or participants to the different conditions); e) compared a multiple contexts condition with a single context condition, and f) reported effect size statistics or, otherwise, enough descriptive data to calculate them. In the cases in which the presence of one or more of these elements was unclear, the authors examined the articles to reach a consensus for inclusion or exclusion. These criteria for study selection also serve to control for bias and/or quality of the individual studies, since they also assess the method and design with which the studies were conducted.

Data Collection

Data from all selected articles were extracted and codified by two independent reviewers (JB and MS). For the codification process, a data extraction sheet was developed by four of the researchers (JB, MS, GM, and ML), according to the follo a) Title; b) Publication date; c) Research group and/or laboratory in which the experiment was performed; d) Experimental subjects (i.e., rodents, humans, etc.); e) Type of sample (e.g., clinical, pre-clinical, etc.); f) Experimental tasks (e.g., fear conditioning, taste aversion, etc.); g) Dependent variables (e.g., number of responses per time unit, latency of response, etc.); and h) Other elements of the experimental design and/or the task: CS, US, number of trials and sessions, number of extinction contexts, type of recovery, and statistical data. Importantly, an "experiment" was defined, according to the eligibility criteria, as any individual comparison between a multiple and a single context condition; thus, a factorial design that compared, for instance, the effect of multiple contexts extinction on both ABA and ABC renewal, would be analyzed as two different experiments instead of one.

Measures and Data Analysis

Statistical analyses were performed using Cohen's d (Cohen, 1988). Values between 0.2 and 0.5 are conventionally considered to indicate a small effect size, between 0.5 and 0.8 a medium effect size, and 0.8 and larger indicate a large effect size. When effect size data was not directly available in the study, it was estimated from descriptive data using the online calculator made available by Wilson (n.d.; available on https://campbellcollaboration.org/research-resources/effect-size-calculator.html). When no descriptive data was provided, effect size statistics were either calculated from descriptive data (means and standard error of the mean or SEM) extracted from the figures using the WebPlotDigitizer online tool (Rohatgi, 2020; available on https://automeris.io/WebPlotDigitizer) or obtained through direct communication with the corresponding author.

Based on the extracted data, an estimation of Cohen's *d* for each critical comparison was performed using the different dependent variables available for each experiment, that is, any measure that provided a direct comparison between performance after extinction in single and multiple contexts. All dependent variables were codified, and their effect size computed for the analyses.

Analyses were then performed based on a multilevel approach using the MetaForest package for R (which includes the commonly used Metafor package; Viechtbauer, 2010; Van Lissa, 2017), which fitted the data into a three-level model. A multilevel meta-analysis is the recommended technique when the effect sizes are dependent and/or nested, and the correlation between variables is unknown (e.g., Fernández-Castilla et al., 2020). This approach is required in this case since experiments in this field usually examine the effect of the manipulation across several dependent variables.

Heterogeneity was examined using Q (Hedges, 1982); a significant Q-value indicates that the studies in the meta-analysis are unlikely to have a common effect size. Alpha was set to .05. An exploratory analysis of relevant categorical moderators was also conducted using the MetaForest package (Van Lissa, 2017), which applies a machine learning approach based on a tree algorithm to the raw effect sizes. A "tree algorithm" divides the data into groups by selecting one moderator, and finding the value that leads to the most homogeneous post-split group. This division is repeated after each split, until a criterion is reached or the number of cases is too small to continue. MetaForest conducts the analysis by drawing a number of bootstrap samples on which a tree grows. Then, all predictions by the trees are averaged, thus selecting the categorical moderators which show a consistent effect on the replications. In this case, the retained moderators had to have an effect on more than 50% of the replications.

Finally, no publication bias analysis was conducted, considering that, to our knowledge, there is yet no reliable technique to assess it within a multilevel meta-analysis (for a thorough analysis, see Fernández-Castilla et al., 2021).

Results

Study Selection

Results of the search can be seen in the following PRISMA flow diagram (Figure 1). A total of 402 studies were found across all sources, of which 101 were first found on Web of Knowledge, 40 on Scopus, 45 on APA PsycInfo, 215 on ProQuest Dissertations & Theses Global and one (consisting of unpublished data) that was received directly from one researcher. After eliminating duplicates, 332 studies remained: 101 obtained on Web of Knowledge, two on Scopus, 14 on APA PsycInfo, 214 on ProQuest Dissertations & Theses Global, and one from a researcher. No studies were found in the references lists. Afterward, these 332 studies were screened by title and abstract; 302 were discarded because no extinction training was included, and the remaining 30 studies were screened by text. Of these 30, four were discarded because effect size statistics were not reported, and it was not possible to extract them from figures or descriptive data; one study was discarded because it did not compare multiple contexts extinction with a single context condition. Thus, the final sample consisted of 25 studies, in English, published between 1998 and 2023, and from which 37 individual experiments (or critical comparisons), with 57 effect sizes, were included in the analyses (see Table 1).

Effect of Extinction in Multiple Contexts on Response Recovery

Effect sizes and the corresponding confidence intervals for each experiment are depicted in Figure 2. The results of the multilevel analysis showed a large aggregate effect size for the 37 experiments and 57 effect sizes, Hedge's $g = 0.92$, CI95%[0.68-1.16]. The Q-value obtained was 257.71, which is higher than the critical chi-square value ($df = 56$), indicating that the sample was heterogeneous. The variance distribution across the three levels of the analysis was 18.82% in Level 1 (participants), 32.74% in Level 2 (within-studies), and 48.42% in Level 3 (between-studies).

Moderator Analysis

The exploratory moderator analysis conducted permutations using the categories of "CS exposure", which divided the studies according to "long", "medium" or "short" exposure times based on the CS duration multiplied by the number of extinction trials ("long" consisted of studies with more than 900 s of total exposure; "medium" between 300 s and 900 s of exposure, and "short" under 300 s; "NA" depicts those studies for which exposure time could not be estimated); "experimental subject" (rodents or humans); "type of sample" (clinical, pre-clinical or non-pathological); "experimental task" (fear conditioning, predictive learning, alcohol reactivity/tolerance or instrumental conditioning), and "recovery type" (ABA renewal, ABC renewal, or ABC renewal plus spontaneous recovery). The results of the analyses (shown in Figure 3) indicated that "CS exposure" had a recursive variable importance of 86%, that is, it had an effect in 86% of the replications; "experimental subject" had a variable importance of 83%, and "type of recovery" of 66%. Both "type of sample" and "experimental task" failed to reach the 50% threshold, with recursive variables importance of 48% and 32%, respectively.

A partial dependence analysis further explored the direction of the effect within each of the retained moderators (Figure 3). The analyses regarding CS exposure showed that on average both short and long exposure times were associated with somewhat greater effect sizes than moderate exposure times; in the case of experimental subjects, experiments conducted with human participants had overall a greater effect size than experiments with rodents. Finally, regarding the type of recovery, experiments that examined ABC renewal plus spontaneous recovery or ABA renewal showed greater effect sizes than those conducted in ABC renewal. It is worth noting, however, that a partial dependence analysis does not conduct hypothesis testing; thus, it does not show whether these differences are significant, only that they are on average numerically different (Van Lissa, 2017).

Discussion

The present multilevel meta-analysis examined the effect of extinction in multiple contexts on response recovery. The findings show that, overall, extinction in multiple contexts has a large effect in reducing response recovery. Heterogeneity, measured with Q, was also significant. The exploratory moderator analysis indicated that the categories with impact on the effect sizes were exposure time to the CS, the experimental subject, and the type of recovery.

The main relevance of the present meta-analysis lies in its multi-level approach, which integrated effect sizes obtained from different dependent variables, across different experimental preparations, and with different experimental subjects. The results show that extinction in multiple contexts has a large effect on response recovery, and a high level of confidence as shown by the confidence interval around the average effect size. Thus, the present results indicate that extinction in multiple contexts might be a useful tool for clinicians in this regard.

These results are however qualified by the heterogeneity of the sample, as shown by Q. The high level of heterogeneity in this case means that there are differences between studies that are not explained by either sampling error or by the assumed distribution of the effect sizes (Cooper, 2017), and thus do not allow the assumption that the effect is similar in every context. In this regard, the moderator analysis suggested several sources for this heterogeneity. The moderators with the largest impact on the effect sizes were experimental subject (with a greater effect in humans compared to rodents); type of recovery (with the effect on ABC renewal being smaller than on both ABA, and ABC renewal plus spontaneous recovery), and exposure time (with a smaller effect in moderate exposure time than in both long and short exposure). The first result is somewhat counterintuitive in that animal studies should allow a more complete experimental control, and less variability in outcomes than human studies. This should in turn lead to more consistent and larger effects. This assumption seems to be unfounded, at least according to the present data, which suggests that human studies are equally useful in detecting large effect sizes.

Regarding the effect of exposure time, at first glance it is consistent with previous evidence suggesting that extinction in multiple contexts with longer extinction or exposure times might be particularly effective in reducing recovery (e.g., Laborda & Miller, 2013; Thomas et al., 2009); however, the present results show that shorter exposure times are also associated with larger effect sizes; that is, both small and large amounts of extinction training in multiple contexts are similarly effective. Why it is the case is not clear; further research might aim at examining this issue more systematically. Finally, the moderator analysis also showed that the effect of extinction in multiple contexts was larger when it was tested on ABC renewal plus Spontaneous Recovery, or on ABA renewal, compared to only ABC Renewal. This is consistent with previous evidence showing that, typically, ABC renewal is weaker than ABA renewal, and sometimes even not observed (e.g., Harris et al., 2000; Neumann, 2006; Üngör & Lachnit, 2006). While at first this result might appear counterintuitive (since the treatment should have a greater impact on weaker types of recovery), it is likely that a stronger recovery makes it easier to observe any decrement since it provides with a stronger comparison baseline; furthermore, from an associative viewpoint it is also likely that there is a larger effect with stronger recovery types as a result of a stronger error correction (e.g., Rescorla & Wagner, 1972). In other words, a stronger renewal should be the result of a stronger recall of the acquisition context compared to extinction; regardless of the associative mechanisms behind it, if extinction in multiple contexts counteracts this recall, the effect should be more pronounced with a stronger renewal than with a smaller one.

No clear recommendations for clinicians are evident from these analyses. From the first one, it can only be surmised that there is a reliable effect of extinction in multiple contexts in human participants across different preparations, and not only in animal studies. The length of exposure time does not offer a clear guideline, since both short and long exposure times are associated with a larger effect; neither does the recovery type, considering that in therapy there is usually no report of the type of recovery being treated, and as such, it is not a relevant factor (e.g., Laborda & Miller, 2012).

A qualitative assessment of the variability within the evidence shows that there is a large diversity of experimental tasks and approaches to the study of extinction and relapse. The present meta-analysis integrated experiments conducted in both rodents and humans, and in several highly different experimental tasks. Although only one of the moderators (experimental subject) can be considered as a procedural variable, the codification of the experiments revealed a high variability within each moderator, and also within studies. For instance, although all experiments in fear conditioning with human participants correspond theoretically to the same procedure, all differ in variables such as the dependent measures used, type and length of CS, number and duration of trials, etc. As a result, several categories were not useful for the analysis because of small sample sizes, which might introduce variability that is not necessarily expressed in the present analysis.

On a conceptual level, one final element to consider is that, empirically speaking, there is no certainty that the present meta-analysis integrated a single effect; it is possible that when we compare, for instance, extinction of fear conditioning and of magazine approach, we are effectively comparing different response systems (e.g., Fanselow, 1994; Fanselow & Wassum, 2016) with different features and mechanisms. Arguably, the main approach to this issue in historical terms has been to consider the several learning tasks promoting a single phenomenon with similar underlying mechanisms, which is reflected in the highly diverse studies integrated into this synthesis. One way to improve future syntheses of the evidence in this field would be to standardize the different experimental tasks and procedures to some extent (e.g., Vervliet, Baeyens, et al., 2013); whenever possible, the different studies using a given learning task should implement similar parameters and procedures. This would probably improve replicability and comparability in the field and make the assessment of the evidence much easier for both theoretical and applied purposes. Such standardization effort might take form, for instance, as a pre-registered set of experiments conducted in different laboratories with similar experimental subjects and parameters, with the aim of assessing the effectiveness of manipulations such as extinction in multiple contexts on recovery and/or relapse in a consistent and probably more definitive manner.

Regarding the limitations of the present synthesis, at least two relevant elements are related to the overall quality and accessibility of the evidence. The first one is the nature of the statistical analyses reported in the literature. Older studies rely almost exclusively on a report of the *p*-value, while estimates of statistical power and reports of effect size are usually absent, although this changes gradually with more recent studies (see e.g., Gunther et al., 1998; Bouton et al., 2006, for examples of the first case, and Krisch et al., 2018, for a more recent and thorough statistical report). This resulted in that most effect sizes had to be estimated based on descriptive data either provided in the studies or extracted from their figures. These estimations show, as depicted in Figure 2, that even in studies that reported a statistically significant effect, confidence intervals often include zero, meaning that there is a high likelihood that in such cases there is no effect, or it is very small. Researchers in this field and others should strive to report comprehensive descriptive data and statistical analyses, including at least effect sizes with their corresponding confidence intervals.

Second, the evidence analyzed in the present synthesis is also limited in the sense that several sources (e.g., conferences or congress abstracts) were not included, and the search was limited to electronic databases, even in the case of gray literature. An effort in contacting each corresponding author for non-published data yielded only one result. The results of the search strategy thus do not include most of the potential unpublished data.

Related to the last point, the lack of publication bias analysis represents a third limitation of the present meta-analysis. To our knowledge, there is no reliable method to analyze selection bias in multi-level meta-analysis. According to Fernández-Castilla et al. (2021), the most used bias analyses (e.g., Sterne & Egger, Funnel plot Test, and Trim and Fill) are overall inadequate, unless several different conditions are met for each technique. For example, their simulations showed that Trim & Fill worked well when the population effect size was moderate to large, there was a high variability among effect sizes, and many effect sizes were included in the analysis. Considering this uncertainty regarding the usefulness of the analysis, and whether or not all assumptions are met by the present sample, publication bias was not assessed; this means that there is no certainty that the sample of this analysis is representative of the evidence.

Overall, the data suggest that extinction in multiple contexts should be an effective manipulation for reducing relapse in exposure therapy. Even if we take the heterogeneity of the sample into account, in all sub-samples the effect size varied from moderate to large, and with reliable confidence intervals; thus, the present results suggest that the effect of multiple contexts extinction appears to be general and should be recommended as a technique for clinicians.

Competing Interests Statement

The authors have no competing interests to declare.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

CRediT Author Contributions

Javier Bustamante: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data Curation, Writing - Original Draft, Writing - Review & Editing, Visualization, Supervision, and Project administration. **Marcela Soto**: Conceptualization, Methodology, Validation, Investigation, Data Curation, Writing - Review & Editing, and Visualization. **Gonzalo Miguez**: Conceptualization, Methodology, Validation, Formal analysis, Data Curation, Writing - Review & Editing, and Visualization. **Vanetza Quezada-Scholz**:

Validation, Writing - Original Draft, and Writing - Review & Editing. **Rocío Angulo**: Validation, Writing - Original Draft, and Writing - Review & Editing. **Mario A. Laborda**: Conceptualization, Methodology, Validation, Data Curation, Writing - Review & Editing, Visualization, and Supervision.

Acknowledgment

We are grateful to the National Agency for Research and Development of Chile (ANID-Chile) for supporting the work of the Research Group in Experimental Psychology and Psychopathology, Department of Psychology, Faculty of Social Sciences, University of Chile. The researchers were partially funded by projects Fondecyt #11200561, #1160132, #11170143, and #1191619, and by ANID PCH/Doctorado Nacional 2019 (#21190913). The authors would like to thank Felipe Alfaro, Francisca Bertin, Esperanza Ibáñez, Luis Pardo, Simón Ramírez, María de los Ángeles Rodríguez, Consuelo San Martín, and Felipe Varas for their comments on earlier versions of this manuscript. Inquiries concerning this research should be addressed to Mario A. Laborda, Department of Psychology, University of Chile; e-mail: mariolaborda@u.uchile.cl.

References

***Articles integrated in the meta-analysis.

Alfaro, F., Navarro, V. M., Laborda, M. A., Betancourt, R., & Miguez, G. (2019). The role of stimuli associated with extinction cues on response recovery [Rol de estímulos asociados a las claves de extinción en la recuperación de respuesta]. *Psykhe, 28*(1), 1-15. https://doi.org/10.7764/psykhe.28.1.1282

Appelbaum, M., Cooper, H., Kline, R. B., Mayo-Wilson, E., Nezu, A., & Rao, S. M. (2018). Journal article reporting standards for quantitative research in psychology: The APA

Publications and Communications Board Task Force Report. *American Psychologist, 73*(1), 3–25. https://doi.org/10.1037/amp0000191

***Bandarian Balooch, S. & Neumann, D. L. (2011). Effects of multiple contexts and context similarity on the renewal of extinguished conditioned behavior in an ABA design with humans. *Learning and Motivation, 42*, 53-63. https://doi.org/10.1016/j.lmot.2010.08.008

***Bandarian-Balooch, S., Neumann, D. L., & Boschen, M. J. (2012). Extinction treatment in multiple contexts attenuates ABC renewal in humans. *Behaviour Research and Therapy, 50*(10), 604-609. https://doi.org/10.1016/j.brat.2012.06.003

***Bandarian-Balooch, S., Neumann, D. L., & Boschen, M. J. (2015). Exposure treatment in multiple contexts attenuates return of fear via renewal in high spider fearful individuals. *Journal of Behavior Therapy and Experimental Psychiatry, 47*, 138-144. https://doi.org/10.1016/j.jbtep.2014.12.006

***Bernal-Gamboa, R., Nieto, J., & Uengoer, M. (2017). Effects of extinction in multiple contexts on renewal of instrumental responses. *Behavioural Processes, 142,* 64-69. https://doi.org/10.1016/j.beproc.2017.06.003

***Bernal-Gamboa, R., Nieto, J., & Gaméz, A. M. (2020). Conducting extinction in multiple contexts attenuates relapse of operant behavior in humans. *Behavioural Processes, 181,* 104261. https://doi.org/10.1016/j.beproc.2020.104261

Betancourt, R., Corada, L., Dominichetti, J., Laborda, M. A., Martínez, G., & Miguez, G. (2008). Effect of extinction in multiple contexts on the renewal of associative ethanol tolerance [Efecto de la extinción en múltiples contextos sobre la renovación de la tolerancia asociativa al etanol]. *Psicothema, 20*(2), 285-289.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta‐analysis*. John Wiley & Sons, Ltd. https://doi.org/10.1002/9780470743386

Bouton, M. E. (1991). A contextual analysis of fear extinction. In P. R. Martin (Ed.), *Handbook of behavior therapy and psychological science: An integrative approach* (pp. 435–453). Pergamon Press, Inc.

Bouton, M. E. (1993). Context, time, and memory retrieval in the interference paradigm of Pavlovian learning. *Psychological Bulletin, 114*, 80-99. https://doi.org/10.1037/0033-2909.114.1.80

Bouton, M. E. (2002). Context, ambiguity, and unlearning: Sources of relapse after behavioral extinction. *Biological Psychiatry, 52,* 976–986. https://doi.org/10.1016/S0006-3223(02)01546-9

Bouton, M. E. & Bolles, R. C. (1979). Contextual control of the extinction of conditioned fear. *Learning and Motivation, 10*, 445-466. https://doi.org/10.1016/0023-9690(79)90057-2

***Bouton, M. E., García-Gutiérrez, A., Zilski, J., & Moody, E. W. (2006). Extinction in multiple contexts does not necessarily make extinction less vulnerable to relapse. *Behaviour Research and Therapy, 44*, 983-994. https://doi.org/10.1016/j.brat.2005.07.007

Bouton, M. E. & King, D. A. (1983). Contextual control of the extinction of conditioned fear: Tests for the associative value of the context. *Journal of Experimental Psychology: Animal Behavior Processes, 9*, 248-265. https://doi.org/10.1037/0097-7403.9.3.248

Bouton, M. E., Woods, A. M., Moody, E. W., Sunsay, C., & García-Gutiérrez, A. (2006). Counteracting the context-dependence of extinction: Relapse and some tests of possible methods of relapse prevention. In M. G. Craske, D. Hermans, & D. Vansteenwegen

(Eds.), *Fear and Learning: Basic Science to Clinical Application.* American Psychological Association.

Bouton, M. E., Maren, S., & McNally, G. P. (2021). Behavioral and neurobiological mechanisms of Pavlovian and instrumental extinction learning. *Physiological Reviews, 101*(2), 611-681. https://doi.org/10.1152/physrev.00016.2020

Bowers, M. E. & Ressler, K. J. (2015). An overview of translationally informed treatments for posttraumatic stress disorder: Animal models of pavlovian fear conditioning to human clinical trials. *Biological Psychiatry, 78*(5), E15-E27. https://doi.org/10.1016/j.biopsych.2015.06.008

Bustamante, J., San Martín, C., Laborda, M. A., & Miguez, G. (2019). An extinction cue does not necessarily prevent response recovery after extinction. *Learning and Motivation, 67*, 101576. https://doi.org/10.1016/j.lmot.2019.101576

***Bustamante, J., Uengoer, M., Thorwart, A., & Lachnit, H. (2016). Extinction in multiple contexts: Effects on the rate of extinction and strength of response recovery. *Learning & Behavior, 44*(3), 283-294. https://doi.org/10.3758/s13420-016-0212-7

Chambless, D. L. & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685-716. https://doi.org/10.1146/annurev.psych.52.1.685.

Chao, W. J. M., Wong, A. V., & McConnell, B. L. (2021). The effects of extinction in multiple spatial contexts on fear recovery in animals and humans: A meta-analysis protocol. PREPRINT (Version 1) available at Research Square https://doi.org/10.21203/rs.3.rs-144660/v1

***Chelonis, J. J., Calton, J. L., Hart, J. A., & Schachtman, T. R. (1999). Attenuation of the renewal effect by extinction in multiple contexts. *Learning and Motivation, 30*, 1-14. https://doi.org/10.1006/lmot.1998.1022

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates, Publishers.

Cooper, H. (2017). *Research synthesis and meta-analysis: A step-by-step approach* (5th ed.). Sage Publications, Inc.

Cooper, H. (2018). *Reporting quantitative research in psychology: How to meet APA style reporting standards* (2nd ed.). American Psychological Association. https://doi.org/10.1037/0000178-000

***Cordero, J. M., López Guerra, V. M., & Bonfante R. A. (2014). Effect of extinction in multiple contexts on the renewal of conditioned compensatory responses to ethanol. [Efecto de la extinción en múltiples contextos sobre la renovación de las respuestas compensatorias condicionadas al etanol]. *Revista Electrónica de Psicología Iztacala, 17*(4), 1622-1647.

Craske, M. G., Hermans, D., & Vervliet, B. (2018). State of the art and future directions for extinction as a translational model for fear and anxiety. *Philosophical Transactions Royal Society B, 373*, 20170025. https://doi.org/10.1098/rstb.2017.0025.

Crombie, K. M., Sartin-Tarm, A., Sellnow, K., Ahrenholtz, R., Lee, S., Matalamaki, M., Adams, T. G., & Cisler, J. M. (2021). Aerobic exercise and consolidation of fear extinction learning among women with posttraumatic stress disorder. *Behaviour Research and Therapy, 142*, 103867. https://doi.org/10.1016/j.brat.2021.103867

Delamater, A. R. (2004). Experimental extinction in Pavlovian conditioning: Behavioural and neuroscience perspectives. *Quarterly Journal of Experimental Psychology Section B*, *57* (2), 97-132. https://doi.org/10.1080/02724990344000097

Delamater, A. R. & Westbrook, R. F. (2014). Psychological and neural mechanisms of experimental extinction: a selective review. *Neurobiology of Learning and Memory*, *108*, 38-51. https://doi.org/10.1016/j.nlm.2013.09.016

***Dunsmoor, J. E., Ahs, F., Zielinski, D. J., & LaBar, K. S. (2014). Extinction in multiple virtual reality contexts diminishes fear reinstatement in humans. *Neurobiology of Learning and Memory, 113*, 157-164. https://doi.org/10.1016/j.nlm.2014.02.010

Fanselow, M. S. (1994). Neural organization of the defensive behavior system responsible for fear. *Psychonomic Bulletin & Review, 1*, 429–438. https://doi.org/10.3758/BF03210947

Fanselow, M. S. & Wassum, K. M. (2016). The origins and organization of vertebrate Pavlovian conditioning. *Cold Spring Harbor Perspectives in Biology*, *8*(1), a021717. https://doi.org/10.1101/cshperspect.a021717

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2020). The application of meta-analytic (multi-level) models with multiple random effects: A systematic review. *Behavior Research Methods*, *52*, 2031-2052. https://doi.org/10.3758/s13428-020-01373-9

Fernández-Castilla, B., Declercq, L., Jamshidi, L., Beretvas, S. N., Onghena, P., & Van den Noortgate, W. (2021). Detecting selection bias in meta-analyses with multiple outcomes: a simulation study. *The Journal of Experimental Education*, *89*(1), 125-144. https://doi.org/10.1080/00220973.2019.1582470

***Glautier, S., Elgueta, T., & Nelson, J. B. (2013). Extinction produces context inhibition and multiple contexts extinction reduces response recovery in human predictive learning. *Learning & Behavior, 41*(4), 341-352. https://doi.org/10.3758/s13420-013-0109-7

***González, V. V., Navarro, V. Miguez, G., Betancourt, R., & Laborda, M. A. (2016). Preventing the recovery of extinguished ethanol tolerance. *Behavioural Processes, 124,* 141-148. https://doi.org/10.1016/j.beproc.2016.01.004

***Gunther, L. M., Denniston, J. C., & Miller, R. R. (1998). Conducting exposure treatment in multiple contexts can prevent relapse. *Behaviour Research and Therapy, 36*, 75-91. https://doi.org/10.1016/s0005-7967(97)10019-5

Harris, J. A., Jones, M. L., Bailey, G. K., & Westbrook, R. F. (2000). Contextual control over conditioned responding in an extinction paradigm. *Journal of Experimental Psychology: Animal Behavior Processes*, *26*, 174–185. https://doi.org/10.1037/0097-7403.26.2.174

Hedges, L. V. (1982). Fitting categorical models to effect sizes from a series of experiments. *Journal of Educational Statistics, 7*(2), 119–137. https://doi.org/10.2307/1164961

***Hermann, A., Stark, R., Müller, E. A., Kruse, O., Wolf, O. T., & Merz, C. J. (2020). Multiple extinction contexts modulate the neural correlates of context-dependent extinction learning and retrieval. *Neurobiology of Learning and Memory, 168*, 107150. https://doi.org/10.1016/j.nlm.2019.107150.

***Krisch, K. A., Bandarian-Balooch, S., & Neumann, D. I. (2018). Effects of extended extinction and multiple extinction contexts on ABA renewal. *Learning and Motivation, 63*, 1-10. https://doi.org/10.1016/j.lmot.2017.11.001

Laborda, M. A., McConnell, B. L., & Miller, R. R. (2011). Behavioral techniques to reduce relapse after exposure therapy: Applications of studies of experimental extinction. In T.

R. Schachtman & S. Reilly (Eds.), *Associative learning and conditioning theory: Human and non-human applications* (pp. 79–103). Oxford University Press.

Laborda, M. A. & Miguez, G. (2015). Exposure therapy. In E. S. Neukrug (Ed.), *Encyclopedia of Theory in Counseling & Psychotherapy* (vol. 1, pp. 385-388). SAGE.

Laborda, M. A. & Miller, R. R. (2011). S-R associations, their extinction, and recovery in an animal model of anxiety: a new associative account of phobias without recall of original trauma. *Behavior Therapy*, *42*(2), 153-169. https://doi.org/10.1016/j.beth.2010.06.002

***Laborda, M. A. & Miller, R. R. (2013). Preventing return of fear in an animal model of anxiety: Additive effects of massive extinction and extinction in multiple contexts. *Behavior Therapy, 44*, 249-261. https://doi.org/10.1016/j.beth.2012.11.001

Laborda, M. A., Polack, C. W., Miguez, G., & Miller, R. R. (2014). Behavioral techniques for attenuating the expression of fear associations in an animal model of anxiety. *Journal of Behavior Therapy and Experimental Psychiatry, 45*(3), 343-350. https://doi.org/10.1016/j.jbtep.2014.02.005

Lipp, O. V., Waters, A. M., Luck, C. C., Ryan, K. M., & Craske, M. G. (2020). Novel approaches for strengthening human fear extinction: The roles of novelty, additional USs, and additional GSs. *Behaviour Research and Therapy, 124*, 103529. https://doi.org/10.1016/j.brat.2019.103529

***MacKillop, J. & Lisman, S. A. (2008). Effects of a context shift and multiple context extinction on reactivity to alcohol cues. *Experimental and Clinical Psychopharmacology, 16*(4), 322-331. https://doi.org/10.1037/a0012686

McConnell, B. L., Miguez, G., & Miller, R. R. (2013). Extinction with multiple excitors. *Learning & Behavior, 41*(2), 119-137. https://doi.org/10.3758/s13420-012-0090-6

McConnell, B. L. & Miller, R. R. (2014). Associative accounts of recovery-from-extinction effects. *Learning and Motivation, 46*, 1–15. https://doi.org/10.1016/j.lmot.2014.01.003

Mellentin, A. I., Skøt, L., Nielsen, B., Schippers, G. M., Nielsen, A. S., Stenager, E., & Juhl, C. (2017). Cue exposure therapy for the treatment of alcohol use disorders: A meta-analytic review. *Clinical Psychology Review, 57*, 195-207. https://doi.org/10.1016/j.cpr.2017.07.006.

Miguez, G., Laborda, M.A., & Miller, R.R. (2014). Enhancement and reduction of associative retroactive cue interference by training in multiple contexts. *Learning & Behavior*, *42*(4), 318-329. https://doi.org/10.3758/s13420-014-0149-7.

Miguez, G., Witnauer, J. E., Laborda, M. A., & Miller, R. R. (2014). Trial spacing during extinction: The role of context–us associations. *Journal of Experimental Psychology: Animal Learning and Cognition, 40*(1), 81-91. https://doi.org/10.1037/a0033203

Miller, R. R. & Laborda, M. A. (2011). Preventing recovery from extinction and relapse: A product of current retrieval cues and memory strengths. *Current Directions in Psychological Science, 20*, 325-329. https://doi.org/10.1177/0963721411418466

Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. *PLoS Medicine, 6*(7), e1000097. https://doi.org/10.1371/journal.pmed.1000097

Nihei, M., Hojo, D., Tanaka, T., & Sawa, K. (2020). A new model for recovery-from-extinction effects in Pavlovian conditioning and exposure therapy. *PsyArXiv*. https://doi.org/10.31234/osf.io/8bjdp

Neumann, D. L. (2006). The effects of physical context changes and multiple extinction contexts on two forms of renewal in a conditioned suppression task with humans. *Learning and Motivation, 37*, 149-175. https://doi.org/10.1016/j.lmot.2005.06.004

Neumann, D. L., Lipp, O. V., & Cory, S. E. (2007). Conducting extinction in multiple contexts does not necessarily attenuate the renewal of shock expectancy in a fear-conditioning procedure with humans. *Behaviour Research and Therapy, 45*, 385-394. https://doi.org/10.1016/j.brat.2006.02.001

***Olatunji, B. O., Tomarken, A., Wentworth, B., & Fritzsche, L. (2017). Effects of exposure in single and multiple contexts on fear renewal: The moderating role of threat-specific and nonspecific emotionality. *Journal of Behavior Therapy and Experimental Psychiatry, 54*, 270-277. https://doi.org/10.1016/j.jbtep.2016.09.004

O'Malley, K. R. & Waters, A. M. (2018). Attention avoidance of the threat conditioned stimulus during extinction increases physiological arousal generalisation and retention. *Behaviour Research and Therapy, 104*, 51-61. https://doi.org/10.1016/j.brat.2018.03.001

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., Shamseer, L., Tetzlaff, J. M., Akl, E. A., Brennan, S. E., Chou, R., Glanville, J. Grimshaw, J. M., Hróbjartsson, A., Lalu, M. M., Li, T., Loder, E. W., Mayo-Wilson, E., McDonald, S., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Parker, Z. J., Waller, G., Gonzalez-Salas Duhne, P., & Dawson, J. (2018). The role of exposure in treatment of anxiety disorders: A meta-analysis. *International Journal of Psychology & Psychological Therapy, 18*, 111-141.

Pavlov, I. P. (1927). *Conditioned reflexes* (G. V. Anrep, Ed. & Trans.). Oxford University Press.

***Pineño, O. & Miller, R. R. (2004). Signaling a change in cue-outcome relation in human associative learning. *Learning & Behavior, 32*(3), 360-375. https://doi.org/10.3758/BF03196034

Pittig, A., van den Berg, L., & Vervliet, B. (2016). The key role of extinction learning in anxiety disorders: Behavioral strategies to enhance exposure-based treatments. *Current Opinion in Psychiatry, 29*(1), 39-47. https://doi.org/10.1097/YCO.0000000000000220.

Polack, C. W., Laborda, M. A., & Miller, R. R. (2013). On the differences in degree of renewal produced by the different renewal designs. *Behavioural Processes, 99*, 112–120. https://doi.org/10.1016/j.beproc.2013.07.006

Quezada, V., Laborda , M. A., Díaz, M., Navarro, V., Mallea, J., Repetto, P., Orellana, G., & Betancourt, R. (2018). Extinction cues do not reduce recovery of extinguished conditioned fear in humans. *International Journal of Psychology and Psychological Therapy, 18*(1) 39-53.

Rachman, S. (1989). The return of fear: Review and prospect. *Clinical Psychology Review, 9*(2), 147-168. https://doi.org/10.1016/0272-7358(89)90025-1.

Rescorla, R. A. (2004). Spontaneous recovery. *Learning & Memory, 11*(5), 501–509. https://doi.org/10.1101/lm.77504

Rohatgi, A. (2020). *WebPlotDigitizer* (Version 4,4) [Computer software]. Retrieved from https://automeris.io/WebPlotDigitizer

San Martín, C., Diaz, F., Cañete, A., Laborda, M. A., & Miguez, G. (2018). Reacquisition of associative tolerance to ethanol: The effect of massive extinction [Readquisición de la tolerancia asociativa al etanol: el efecto de la extinción masiva]. *Avances en Psicología Latinoamericana, 36*(2), 419-429. https://doi.org/10.12804/revistas.urosario.edu.co/apl/a.5727

Scheveneels, S., Boddez, Y., Vervliet, B., & Hermans, D. (2019). Modeling hierarchical versus random exposure schedules in Pavlovian fear extinction: No evidence for differential fear outcomes. *Behavior Therapy, 55*(5), 967-977. https://doi.org/10.1016/j.beth.2019.03.001

***Shiban, Y., Pauli, P., & Mühlberger, A. (2013). Effects of multiple contexts exposure on renewal in spider phobia. *Behaviour Research and Therapy, 51*, 68-74. https://doi.org/10.1016/j.brat.2012.10.007

***Shiban, Y., Schelhorn, I., Pauli, P., & Mühlberger, A. (2015). Effects of combined multiple contexts and multiple stimuli exposure in spider phobia: A randomized clinical trial in virtual reality. *Behaviour Research and Therapy, 71*, 45-53. https://doi.org/10.1016/j.brat.2015.05.014

Smith, S. M. (1982). Enhancement of recall using multiple environmental contexts during learning. *Memory & Cognition, 10*, 405–412. https://doi.org/10.3758/BF03197642

Smith, S. M. (1985). Effects of number of study environments and learning instructions on free-recall clustering and accuracy. *Bulletin of the Psychonomic Society, 23*, 440–442. https://doi.org/10.3758/BF03329846

Smith, S. M., Glenberg, A. M., & Bjork, R. A. (1978). Environmental context and human memory. *Memory & Cognition, 6*, 342-353. https://doi.org/10.3758/BF03197465

Spix, M., Lommen, M. J. J., & Boddez, Y. (2021). Deleting "fear" from "fear extinction": Estimating the individual extinction rate via non-aversive conditioning. *Behaviour Research and Therapy, 142,* 103869. https://doi.org/10.1016/j.brat.2021.103869

Thomas, B. L., Vurbic, D., & Novak, C. (2009). Extensive extinction in multiple contexts eliminates the renewal of conditioned fear in rats. *Learning and Motivation, 40*, 147-159. https://doi.org/10.1016/j.lmot.2008.10.002

Üngör, M., & Lachnit, H. (2006). Contextual control in discrimination reversal learning. *Journal of Experimental Psychology: Animal Behavior Processes*, *32*, 441–453. https://doi.org/10.1037/0097-7403.32.4.441

Uengoer, M., Lissek, S., Tegenthoff, M., Manahan-Vaughan, D., & Lachnit, H. (2020). Principles of extinction learning of nonaversive experience. *Neuroforum, 26*(3), 151-159. https://doi.org/10.1515/nf-2020-0013

Van den Noortgate, W., López-López, J. A., Marín-Martínez, F., & Sánchez-Meca, J. (2015). Meta-analysis of multiple outcomes: A multilevel approach. *Behavior Research Methods*, *47*, 1274-1294. https://doi.org/10.3758/s13428-014-0527-2

Van Lissa, C. J. (2017). MetaForest: exploring heterogeneity in meta-analysis using random forests. Open Science Framework, September. https://doi.org/10.17605/OSF.IO/KHJGB

***Vansteenwegen, D., Vervliet, B., Iberico, C., Baeyens, F., Van den Bergh, O., Hermans, D. (2007). The repeated confrontation with videotapes of spiders in multiple contexts attenuates renewal of fear in spider-anxious students. *Behaviour Research and Therapy, 45*, 1169-1179. https://doi.org/10.1016/j.brat.2006.08.023

Vervliet, B., Baeyens, F., Van den Bergh, O., & Hermans, D. (2013). Extinction, generalization, and return of fear: A critical review of renewal research in humans. *Biological Psychology, 92*(1), 51–58. https://doi.org/10.1016/j.biopsycho.2012.01.006

Vervliet, B., Craske, M. G., & Hermans, D. (2013). Fear extinction and relapse: State of the art. *Annual Review of Clinical Psychology, 9*, 215–248. https://doi.org/10.1146/annurev-clinpsy-050212-185542

***Viar-Paxton, M. A. & Olatunji, B. O. (2012). Context effects on habituation to disgust-relevant stimuli. *Behavior Modification, 36*(5), 705-722. https://doi.org/10.1177/0145445512446189

Viechtbauer, W. (2010). Conducting Meta-Analyses in R with the metafor Package. *Journal of Statistical Software*, *36*(3), 1-48. https://doi.org/10.18637/jss.v036.i03

Vurbic, D. & Bouton, M. E. (2014). A contemporary behavioral perspective on extinction. In F. K. McSweeney & E. S. Murphy (Eds.), *The Wiley Blackwell handbook of operant and classical conditioning* (pp. 53–76). Wiley Blackwell. https://doi.org/10.1002/9781118468135.ch3

Waters, A. M., Ryan, K. M., Luck, C. C., Craske, M. G., & Lipp, O. V. (2021). The effects of presenting additional stimuli resembling the CS+ during extinction on extinction retention and generalisation to novel stimuli. *Behaviour Research and Therapy, 114*, 103921. https://doi.org/10.1016/j.brat.2021.103921

Weisman, J. S. & Rodebaugh, T. L. (2018). Exposure therapy augmentation: A review and extension of techniques informed by an inhibitory learning approach. *Clinical Psychology Review, 59*, 41-51. https://doi.org/10.1016/j.cpr.2017.10.010

Wilson, D. B. (n.d.). Practical meta-analysis effect size calculator [Online calculator]. Retrieved from https://campbellcollaboration.org/research-resources/effect-size-calculator.html

***Wong, V. A., Chao, W. J. M., & McConnell, B. L. (2023). Acquisition and extinction across multiple virtual reality contexts. *Learning and Motivation*, *81*, 101857. https://doi.org/10.1016/j.lmot.2022.101857

Zbozinek, T. D. & Craske, M. G. (2018). Pavlovian extinction of fear with the original conditional stimulus, a generalization stimulus, or multiple generalization stimuli. *Behaviour Research and Therapy, 107*, 64-75. https://doi.org/10.1016/j.brat.2018.05.009

Figure 1

PRISMA Flow Diagram for Articles Search and Selection

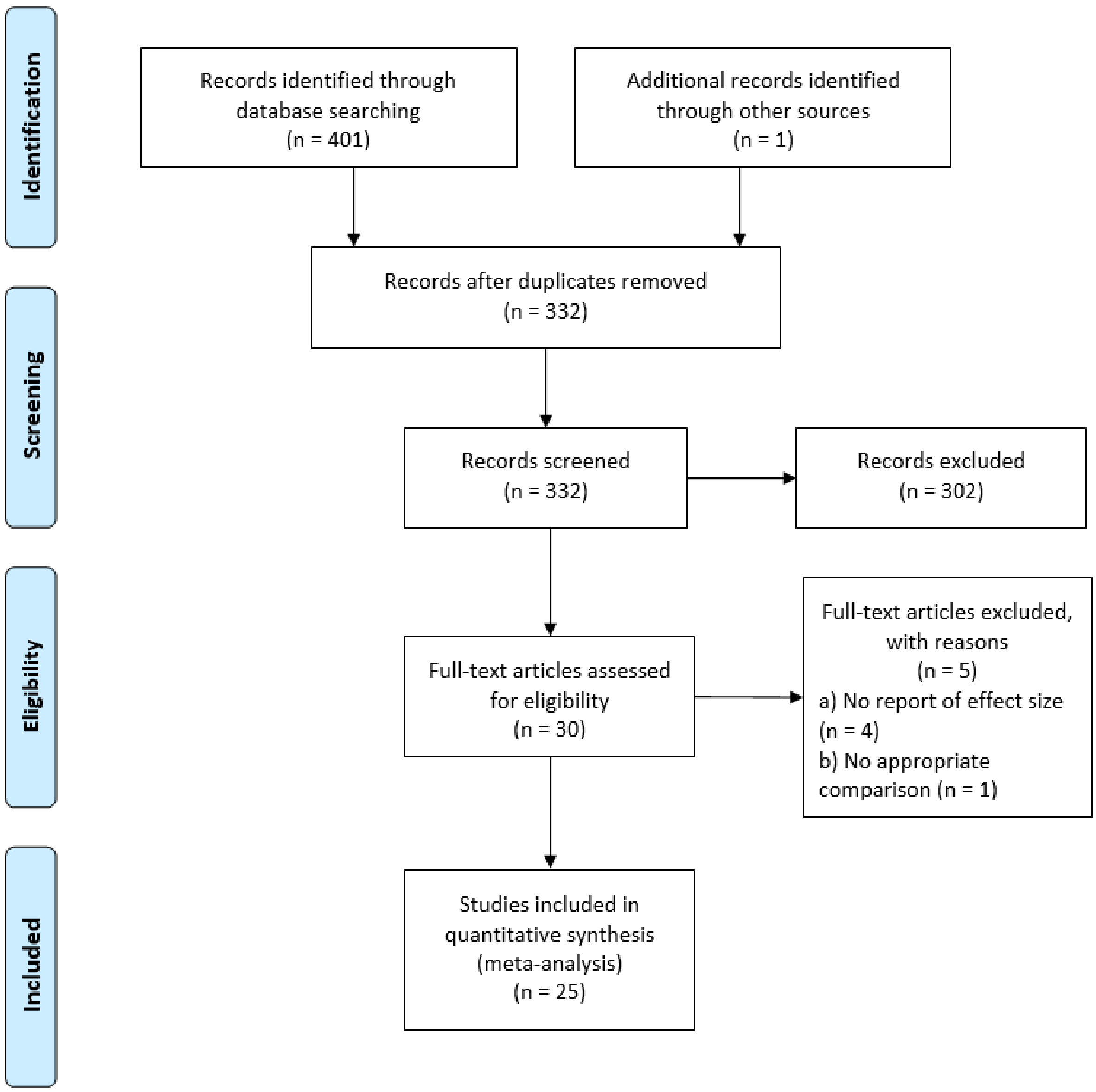
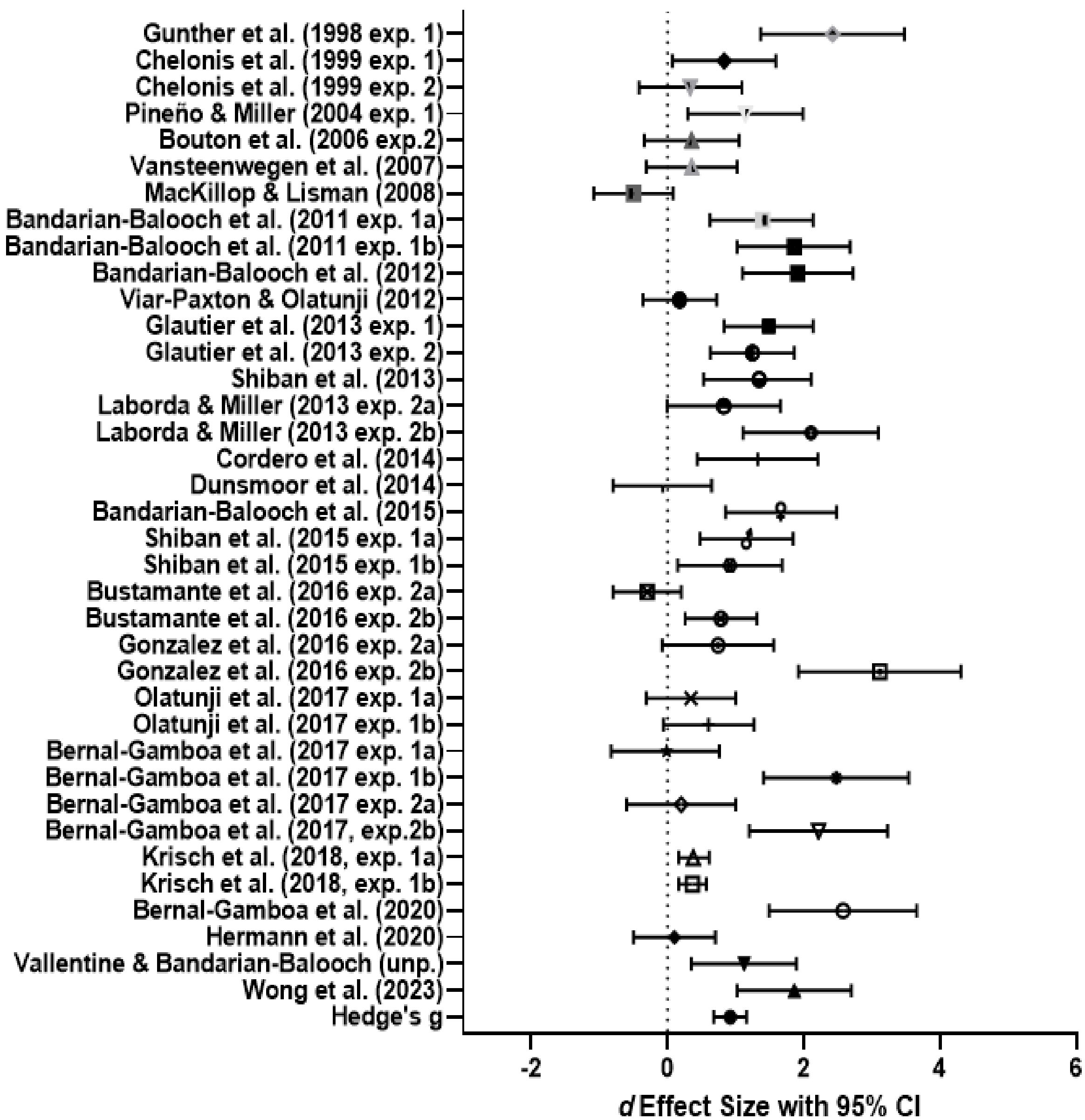


Figure 2

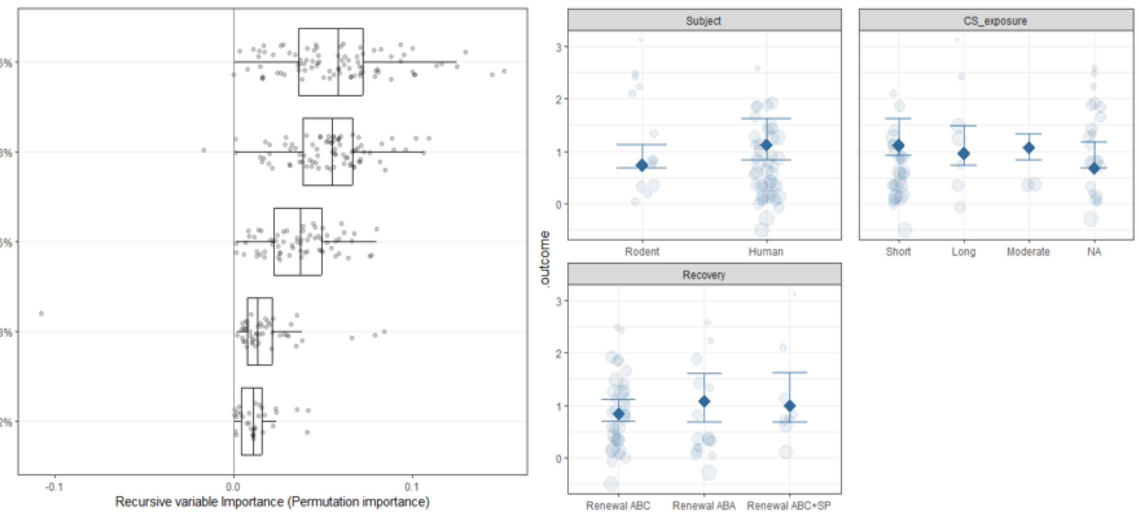
Estimated effect sizes and confidence intervals of each critical comparison



Note. Lines indicate the effect sizes and the corresponding 95% confidence intervals. The last line represents the integrated effect size. Effect sizes for each dependent variable are reported separately in the same order as in Table 1.

Figure 3

Moderator analysis (variable importance and partial dependence)



Note. Points indicate individual effect sizes. Brackets indicate the 95% percentile interval of the predictions of individual trees of the model.

Table 1

Characteristics of the studies included in the meta-analysis

Articles	Critical Comparisons	Subjects or participants	Sex	Experimental Task	Type of Recovery	Dependent Variables	Total Exposure Time
Gunther et al. (1998)	1	Rodent (Sprague-Dawley)	Both	Fear conditioning	Renewal ABC	Conditioned suppression (lick suppression)	Long
Chelonis et al. (1999)	2	Rodent (Sprague-Dawley)	Both	Conditioned taste aversion	Renewal ABA	Conditioned suppression (of water)	-
Pineño & Miller (2004)	1	Human (General)	Both	Predictive learning	Renewal ABC	Conditioned suppression (button suppression)	Short
Bouton et al. (2006)	1	Rodent (Wistar)	Female	Fear conditioning	Renewal ABA	Conditioned suppression (button suppression)	Moderate
Vansteenwegen et al. (2007)	1	Human (Preclinical)	Both	Fear conditioning	Renewal ABC	Self-report (fear); Skin conductance; valence report; disgust report	Moderate
McKillop & Lisman (2008)	1	Human (General)	Both	Alcohol reactivity	Renewal ABC	Self-report (urge to drink); salivation	Long
Bandarian-Balooch & Neumann (2011)	2	Human (General)	Both	Fear conditioning	Renewal ABA	Self-report (expectancy of the US); latency	-
Bandarian-Balooch et al. (2012)	1	Human (General)	Both	Fear conditioning	Renewal ABC	Self-report (expectancy of the US); startle blink	-
Viar-Paxton & Olatunji (2012)	1	Human (General)	Both	Conditioned disgust	Renewal ABC	Self-report (distress); skin conductance; BAT	Short
Glautier et al. (2013)	2	Human (General)	Both	Predictive learning	Renewal ABC	Self-report (prediction)	Short
Laborda & Miller (2013)	2	Rodent (Sprague-Dawley)	Both	Fear conditioning	Renewal ABC+ Spontaneous Recovery	Conditioned suppression (lick suppression)	Long
Shiban et al. (2013)	1	Human (Clinical)	Male	Fear conditioning	Renewal ABC	Self-report (expectancy of the US); skin conductance; BAT	Short

Cordero et al. (2014)	1	Rodent (NMRI mouse)	Male	Conditioned ethanol tolerance	Renewal ABA	Ataxic Response	-
Dunsmoor et al. (2014)	1	Human (General)	Both	Fear conditioning	Renewal ABC + Spontaneous Recovery	Fear potentiated startle	Short
Bandarian-Balooch et al. (2015)	1	Human (Preclinical)	Both	Fear conditioning	Renewal ABC	Self-report (fear); BAT; heart rate	-
Shiban et al. (2015)	2	Human (Clinical)	Both	Fear conditioning	Renewal ABA	Self-report (fear); skin conductance; BAT	Short
Bustamante et al. (2016)	2	Human (General)	Both	Predictive learning	Renewal ABA & renewal ABC	Self-report (expectancy of the US)	-
González et al. (2016)	2	Rodent (Sprague-Dawley)	Male	Conditioned ethanol tolerance	Renewal ABC	Ataxic Response	Long
Bernal-Gamboa et al. (2017)	4	Rodent (Wistar)	Female	Instrumental conditioning	Renewal ABA (3) & renewal ABC (1)	Responses per minute	-
Olatunji et al. (2017)	2	Human (Preclinical)	Both	Fear conditioning	Renewal ABC	Self-report (anxiety)	Moderate
Krisch et al. (2018)	2	Human (General)	Both	Fear conditioning	Renewal ABA	Self-report (expectancy of the US)	Short
Hermann et al. (2020)	1	Human (General)	Male	Fear conditioning	Renewal ABC	Skin conductance; neural activation	Short
Bernal-Gamboa et al. (2020)	1	Human (General)	Both	Instrumental conditioning	Renewal ABA	Responses per minute	-
Vallentine & Bandarian-Balooch (n.d.)	1	Human (Preclinical)	Both	Fear conditioning	Renewal ABC + Spontaneous Recovery	Self-report (fear); BAT	-
Wong et al. (2023)	1	Human (General)	Both	Predictive learning	Renewal ABC	Sel-report (expectancy of the US)	-

Note: Relevant features of the included studies are presented.