**Supplementary Procedure**

The purpose of this document is to highlight several ways in which readers can assess the fidelity of our work as it is described in the manuscript.

**Preregistration**

Each study was preregistered. Permanent uneditable links to these preregistrations can be found at the below links:

Experiment 1: <https://osf.io/nkr9g/>

Experiment 2: https://osf.io/nkr9g/

Experiment 3: https://osf.io/nkr9g/

Experiment 4: https://osf.io/nkr9g/

Experiment 5: https://osf.io/3svcg/

Experiment 6: https://osf.io/3svcg/

Experiment 7: <https://osf.io/3svcg/> and https://osf.io/nkr9g/

We made several checks to ensure that our preregistrations matched up with the manuscript. This highlighted a number of ways in which the two diverged. Recent research has found that such preregistration-manuscript incongruences are common and should be explicated (e.g., <https://www.sciencemag.org/news/2019/05/solution-psychology-s-reproducibility-problem-just-failed-its-first-test>). That is, aesthetic standards for manuscripts should not trump precision (see Giner-Sorolla, ‎2012).

**Deviations from pre-registration**.

1. During the review process a number of reviewers commented that the terminology used in our original pre-registration documents was difficult to understand and made the paper inaccessible. Therefore in the revised manuscript (and updated OSF documents: see https://osf.io/u6vtz/) we use a consistent set of terminology throughout the paper (i.e., positive source stimulus (S1), neutral outcome stimulus (O1), neutral target stimulus (T1), negative source stimulus (S2), neutral outcome stimulus (O2), neutral target stimulus (T2)). We believe this new terminology will make the manuscript easier to follow and our work easier to evaluate.
2. Our final analyses deviated from what we stated in the original pre-registration in several ways:
   1. We originally planned to include Stimulus (T1 vs. T2 or O1 .vs O2) as a within subjects factor in our models. However, we later realized that we were primarily interested in an evaluative learning effect itself, rather than the interaction between evaluative learning and stimulus type. Thus we dropped this factor from the model. That is, we first created a difference score (T1 – T2) or (O1-O2) and then submitted these difference scores to either t-tests (to determine if the IR or OEC effect differed from zero) or a one-way ANOVA (with Training Condition as a between subjects factor) to determine if IR or OEC effects differed as a function of the type of training people received. This served to simplify our models and provide additional power.
   2. We stated that we would use the D6 or D4 algorithms when computing IAT scores. Instead we used the D2 algorithm.
   3. We stated that we would engage in a second round of analyses were we would exclude participants from the paper if they failed the memory questions or other exploratory questions. We originally did so in our initial submission of the paper to a journal. However, the reviewers indicated that this information was too much and asked that we omit it. The current version of the manuscript has omitted these analyses and only reports the primary analyses (i.e., impact of training condition on evaluation).
3. Study 7 in the final manuscript was created by combining the extinction and control condition of Study 1C (see https://osf.io/nkr9g/) and the counter conditioning phase of Study 1C (see https://osf.io/3svcg/). We decided to report the results in this way as doing so would allow us to directly compare the impact of extinction and counterconditioning on IR and OEC effects (rather than make those comparisons between studies). These two conditions also differed from the others in this project (i.e., in that they involved an important procedural change: use of mouse instead of keyboard with onscreen presentation of randomly switching response options across trials).

Other previously-detected and rectified incongruences between the manuscript and the preregistration include:

1. The addition of non-pre-registered meta-analyses of Experiments 1-7.
2. Changes to the updated OSF files to provide a direct record of what is now reported in the manuscript. Those looking to compare the original (<https://osf.io/3svcg/> and <https://osf.io/nkr9g/>) and updated OSF files (<https://osf.io/u6vtz/>) can do so via these URLs.

**Power**

Power analyses were conducted for all studies. However, our sample sizes are smaller than our stated data collection stopping rules due to loss of participants via exclusion criteria. This is one of the reasons why we decided to include a meta-analysis – to assess our primary questions with a much larger sample than obtained in any one study.

**Open materials**

All materials to run the studies are openly available via the URLs outlined in the pre-registration section (see above).

Should it be the case that any aspect of our written description of a procedure was not sufficient in some regard, a reader can therefore easily inspect or run this code to know what was actually done with great certainty. This demonstrates the *methodological reproducibility* of the work and the *potential for trivial direct replication* by others.

For all studies, code for data processing and analyses is also available. This demonstrates the *computational reproducibility* of the work.

**Open data**

We have made not only the processed data openly available, but also the raw data. This has many benefits, including allowing others to check the *computational reproducibility* of our results (as noted above), but also their *robustness to alternative analytic approaches*. Moreover, this *enables data reuse* by others for novel purposes we have not foreseen, potentially conserving scarce research resources or allowing others to conduct meta-analyse using raw data.