**Experiment 1**

**IAT**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed an automatic relative preference for CS1 over CS2 on the IAT (*M* = 0.37, *SD* = 0.46), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -0.23, *SD* = 0.45), *t*(98.12) = 6.63, *p* < .001, Cohen’s *d* = 1.31, 95% CI = [0.88, 1.74], BF10 > 5 million[[1]](#footnote-1).

**Self-report ratings**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed a relative preference on the self-report ratings for CS1 over CS2 (*M* = 3.33, *SD* = 4.60), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -4.15, *SD* = 4.48), *t*(98.32) = 8.33, *p* < .001, Cohen’s *d* = 1.65, 95% CI = [1.20, 2.10], BF10 >15 billion.

**Behavioural intentions**

Data from the behavioural intentions question were entered into a multinomial logistic regression with CSI as the reference category. Only results from the CS1-CS2 comparison are relevant to the hypothesis and will be reported here. Results demonstrated that participants responses of CS1 relative to CS2 differed between the two conditions, congruent with training, OR = 13.66, 95% CI = [3.56, 52.44], *p* < .001.

**Experiment 2**

**IAT**

When CS2 eventually shared a color with positive words and CS1 shared a color with negative words, participants showed an automatic relative preference for CS2 over CS1 on the IAT (*M* = -0.26, *SD* = 0.54), whereas when CS2 eventually shared a color with negative words and CS1 shared a color with positive words, they demonstrated a relative preference for CS1 over CS2 (*M* = -0.12, *SD* = 0.60), *t*(100.85) = -1.18, *p* = .24, Cohen’s *d* = -0.23, 95% CI = [-0.62, 0.16], BF10 = 0.38. Note that this result is in the opposite direction to all other experiments.

**Self-report ratings**

Regardless of whether CS1 eventually shared a color with positive or negative words, participants showed a relative preference on the self-report ratings for CS2 over CS1 (CS1-positive condition: *M* = -2.80, *SD* = 5.33; CS1-negative condition: *M* = -1.58, *SD* = 6.03), *t*(100.98) = -1.09, *p* = .28, Cohen’s *d* = -0.21, 95% CI = [-0.61, 0.18], BF10 = 0.35.

Note that this null result is different to all other experiments.

**Behavioural intentions**

Data from the behavioural intentions question were entered into a multinomial logistic regression with CSI as the reference category. Only results from the CS1-CS2 comparison are relevant to the hypothesis and will be reported here. Results demonstrated that participants responses of CS1 relative to CS2 differed between the two conditions, but in the opposite direction to that predicted, OR = 0.22, 95% CI = [0.07, 0.66], *p* = .007. Note that this result is in the opposite direction to all other experiments.

**Experiment 3**

**IAT**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed an automatic relative preference for CS1 over CS2 on the IAT (*M* = 0.21, *SD* = 0.46), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -0.15, *SD* = 0.59), *t*(93.42) = 3.29, *p* = .001, Cohen’s *d* = 0.66, 95% CI = [0.25, 1.08], BF10 = 20.10.

**Self-report ratings**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed a relative preference on the self-report ratings for CS1 over CS2 (*M* = 2.92, *SD* = 5.25), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -2.85, *SD* = 5.02), *t*(92.94) = 5.52, *p* < .001, Cohen’s *d* = 1.13, 95% CI = [0.69, 1.56], BF10 > 45 thousand.

**Behavioural intentions**

Data from the behavioural intentions question were entered into a multinomial logistic regression with CSI as the reference category. Only results from the CS1-CS2 comparison are relevant to the hypothesis and will be reported here. Results demonstrated that participants responses of CS1 relative to CS2 differed between the two conditions, congruent with training, OR = 6.94, 95% CI = [2.03, 23.77], *p* = .002.

**Experiment 4**

**IAT**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed an automatic relative preference for CS1 over CS2 on the IAT (*M* = 0.14, *SD* = 0.46), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -0.12, *SD* = 0.46), *t*(168.75) = 3.79, *p* < .001, Cohen’s *d* = 0.57, 95% CI = [0.27, 0.87], BF10 = 109.

**Self-report ratings**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed a relative preference on the self-report ratings for CS1 over CS2 (*M* = 2.34, *SD* = 4.12), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -2.38, *SD* = 4.09), *t*(169.77) = 7.66, *p* < .001, Cohen’s *d* = 1.15, 95% CI = [0.83, 1.47], BF10 > 5 billion.

**Behavioural intentions**

Data from the behavioural intentions question were entered into a multinomial logistic regression with CSI as the reference category. Only results from the CS1-CS2 comparison are relevant to the hypothesis and will be reported here. Results demonstrated that participants responses of CS1 relative to CS2 differed between the two conditions, congruent with training, OR = 5.00, 95% CI = [1.91, 13.06], *p* = .001.

**Experiment 5**

**IAT**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed an automatic relative preference for CS1 over CS2 on the IAT (*M* = 0.16, *SD* = 0.48), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -0.18, *SD* = 0.45), *t*(184.64) = 4.98, *p* < .001, Cohen’s *d* = 0.73, 95% CI = [0.43, 1.02], BF10 > 10 thousand.

**Self-report ratings**

When CS1 eventually shared a color with positive words and CS2 shared a color with negative words, participants showed a relative preference on the self-report ratings for CS1 over CS2 (*M* = 3.57, *SD* = 4.99), whereas when CS1 eventually shared a color with negative words and CS2 shared a color with positive words, they demonstrated a relative preference for CS2 over CS1 (*M* = -2.18, *SD* = 4.19), *t*(179.22) = 8.54, *p* < .001, Cohen’s *d* = 1.25, 95% CI = [0.93, 1.56], BF10 > 1 trillion.

**Behavioural intentions**

Data from the behavioural intentions question were entered into a multinomial logistic regression with CSI as the reference category. Only results from the CS1-CS2 comparison are relevant to the hypothesis and will be reported here. Results demonstrated that participants responses of CS1 relative to CS2 differed between the two conditions, congruent with training, OR = 7.09, 95% CI = [2.92, 17.21], *p* < .001.

**Meta analyses**

Random effects meta analyses were fitted using the metafor R package (Viechtbauer, 2010) and the maximum likelihood estimator function. A separate meta analysis was fitted for each outcome variable (IAT, self reported ratings, behavioural intentions). Although multivariate meta analysis was a possible alternative, previous work has traditionally assessed these dependent variables separately.

**IAT**

Results demonstrated a significant effect of medium size (Cohen, 1988): *k* = 5, Cohen’s *d* = 0.61, 95% CI = [0.13, 1.08], 95% CR = [-0.49, 1.70], *p* = .012. However, results were found to contain a high degree of heterogeneity, Q(df = 4) = 28.46, *p* = 0.012, τ2 = 0.25, *I*2 = 88.34, *H*2 = 8.57. As such, the presence of outlier experiments was assessed using metrics of excessive influence on the meta analyzed effect size and/or excessive influence on heterogeneity (τ2) via leave-one-out analyses. The consistency of results across all three outcome variables (i.e., IAT, self-reports, and behavioural intentions; see below) was assessed when determining outliers. Assessment of all three outcome variables uniformly suggested that Experiment 2 was an outlier on the basis of undue influence on both the meta analyzed effect size and heterogeneity. As such, Experiment 2 was excluded and a second meta analysis was refit. After exclusions, the meta analyzed effect size was still found to be significant and was now of large size, *k* = 4, Cohen’s *d* = 0.80, 95% CI = [0.49, 1.10], 95% CR = [0.22, 1.37], *p* < .001, and with lower heterogeneity, Q(df = 3) = 8.08, *p* = < .001, τ2 = 0.06, *I*2 = 65.47, *H*2 = 2.90.

**Self-report ratings**

Results demonstrated that the meta analytic effect was significant and of large size, k = *5*, Cohen’s *d* = 0.99, 95% CI = [0.38, 1.60], 95% CR = [-0.46, 2.44], *p* = .002, but with a high degree of heterogeneity: Q(df = 4) = 48.03, *p* = 0.002, τ2 = 0.45, *I*2 = 92.53, *H*2 = 13.39. After excluding Experiment 2 as an outlier, the meta analyzed effect size was still significant but was now of very large size (Sawilowsky, 2009), *k* = 4, Cohen’s *d* = 1.26, 95% CI = [1.08, 1.44], 95% CR = [1.08, 1.44], *p* < .001, and with negligible heterogeneity, Q(df = 3) = 3.66, *p* = < .001, τ2 = 0, *I*2 = 0.01, *H*2 = 1.00.

**Behavioural intentions**

Results demonstrated a non-significant meta analytic effect of medium size (Chen, Cohen, & Chen, 2010), *k* = 5, OR = 3.71, 95% CI = [0.90, 15.33], 95% CR = [0.14, 99.48], *p* = .07, and with a high degree of heterogeneity: Q(df = 4) = 32.06, *p* = 0.07, τ2 = 2.29, *I*2 = 88.29, *H*2 = 8.54. After excluding Experiment 2 as an outlier, the meta analyzed effect size was significant and of large size, *k* = 4, OR = 7.03, 95% CI = [4.14, 11.94], 95% CR = [4.14, 11.94], *p* < .001, and with negligible heterogeneity, Q(df = 3) = 1.42, *p* = < .001, τ2 = 0, *I*2 = 0.00, *H*2 = 1.00.

**Robustness tests**

Following our preregistered analytic plan to examine the role of the exploratory awareness variables, several subsets of participants were excluded and meta analytic models were again refit. This served to assess the robustness of the learning effect to including only participants who (a) were contingency aware, (b) were not demand compliant, (c) were not hypothesis aware, and (d) were not influence aware. These robustness tests were run subsequent to excluding Experiment 2 as an outlier. For all dependent variables (IAT, self-reports, behavioural intentions), and across all the above awareness types, meta analytic learning effects were found to be significant (all *p*s <= .001). This was with one exception: differences in behavioural intentions between the groups were not robust to excluding individuals who were influence aware (*p* = .09).

1. All Bayes Factors employed a default prior employed within the BayesFactor R package: a Cauchy distribution with scaling factor 0.707 placed on the effect size. [↑](#footnote-ref-1)