Supplementary Materials

Feature Repetition as Feature (Dis)Similarity Between Features at the Cued Position in the Cue Display and the Target Features

Pure onset cue blocks

RTs. Validity and relevant repetition interacted, with $F(1, 19) = 22.85$, $p < .001$, $\eta_p^2 = .55$. Compatibility effects were observed in valid trials, 23 ms, $t(19) = 4.07$, $p < .001$, $d = 0.25$, but not in invalid trials ($p = .17$).

ERs. No main effects or interactions were significant.

Mixed-cue blocks

RTs. Validity interacted with irrelevant repetition, $F(1, 19) = 7.15$, $p = .015$, $\eta_p^2 = .27$. Significant compatibility effects arose in valid trials, 11 ms, $t(19) = 2.34$, $p = .03$, $d = 0.11$, but not in invalid trials ($p = .63$).

ERs. Relevant and irrelevant repetition interacted, $F(1, 19) = 6.13$, $p = .023$, $\eta_p^2 = .24$. Relevant repetition effects arose when the irrelevant feature changed (repetition: 3.5% vs. change: 4.7%), $t(19) = 3.01$, $p < .01$, $d = 0.58$. The three-way interaction between validity, relevant and irrelevant repetition was significant, $F(1, 19) = 5.88$, $p = .025$, $\eta_p^2 = .24$. Significant relevant-repetition effects emerged in valid trials when the irrelevant feature changed (repetition: 2.8% vs. change: 5%), $t(19) = 2.46$, $p = .02$, $d = 0.68$. All other ps > .15

The results are again neither fully in line with any of the expected influences under the perspective of the TEC, nor with the assumption that any of the other features (here: of the cue) besides color contributed to object-file updating costs.

DOI 10.1027/1618-3169/10.1027/1618-3169/a000511

Effects of experiment version in Experiment 2

Here we report effects in Experiment 2 that involved the between-participants factor experiment version (lab-based vs. online). We do that for two reasons: Firstly, we wanted to keep the results in the Results section of Experiment 2 brief and comparable to the results in Experiment 1. Secondly, the choice to test participants in an online study was a practical, not a theoretical one. Hence, sample sizes of both experiment versions were not determined based on power considerations that involved the comparison of the two versions. While it is possible that data from the online version (as in Experiment 1) includes variance stemming from luminance differences of the stimulus colors, Experiment 2 was not designed to address this question.

Compatibility as Letter- and Orientation Similarity at the Targets' Position

Response Times. A significant main effect of experiment version was found (44 ms difference: online version: 637 ms; vs. lab-based version: 681 ms), $F(1, 30) = 5.78$, $p = .023$, $\eta_p^2 = .16$.

Experiment version interacted with search difficulty, $F(1, 30) = 13.1$, $p = .001$, $\eta_p^2 = .30$. Participants were 82 ms faster in the online version (679 ms) than in the lab-based version (762 ms) of the difficult search condition $t(30) = 3.2$, $p < .01$, $d = 1.17$. No such difference was found in the easy search condition ($p = .66$).

Experiment version entered a four-way interaction with validity, relevant, and irrelevant repetition, $F(1, 30) = 4.74$, $p = .038$, $\eta_p^2 = .14$. Relevant repetition only sped up RTs in the online version of the experiment under valid and invalid conditions when the irrelevant feature changed; valid condition: 13 ms (relevant repetition: 611 ms vs. relevant change: 624

DOI 10.1027/1618-3169/10.1027/1618-3169/a000511

ms), $t(18) = 3.54$, $p < .01$, $d = 0.29$; invalid condition: 15 ms (relevant repetition: 645 ms vs. relevant change: 660 ms), $t(18) = 7.14$, $p < .001$, $d = 0.28$. In the online version of the experiment, relevant repetition trials *slowed* RTs in invalid trials when the irrelevant feature repeated, −14 ms (relevant repetition: 648 ms vs. relevant change: 634 ms), $t(18) = -6.66$, $p = .001$, $d = -0.29$.

Error Rates. Experiment version interacted with search difficulty, $F(1, 30) = 12.58$, $p = .038$, $\eta_p^2 = .14$. Participants committed more errors in the lab-based version than in the online version of the difficult search condition, 7.2% versus 4.3%, $t(30) = 2.52$, $p = .02$, $d = 0.92$. Again, no difference was found in the easy search condition ($p = .42$).

Experiment version interacted with cue match and irrelevant feature repetition, $F(1, 30) = 6.07$, $p = .02$, $\eta_p^2 = .17$, entered a four-way interaction with validity, search difficulty, and relevant repetition, $F(1, 30) = 5.50$, $p = .03$, $\eta_p^2 = .15$. Finally, Experiment entered two five-way interactions with validity, cue match, relevant repetition, and search difficulty, $F(1, 30) = 6.52$, $p = .016$, $\eta_p^2 = .18$, and with validity, relevant repetition, irrelevant repetition, and search difficulty, $F(1, 30) = 5.49$, $p = .026$, $\eta_p^2 = .15$. Regarding the first five-way interaction, we first looked at error rates from the lab-based version of the experiment. Under easy search conditions, conventional response-compatibility effects were only found for valid matching cues (relevant-repetition: 1.1% vs. relevant-change: 3%), $t(12) = 2.9$, $p = .01$, $d = 0.91$. Under difficult search conditions, response-compatibility effects were only found with invalid non-matching cues (relevant-repetition: 7.2% vs. relevant-change: 9.2%), $t(12) = 2.52$, $p = .03$, $d = 0.48$. Looking at the data from the online version of the experiment, response-compatibility effects under easy search conditions were only found for invalid matching cues (relevant-repetition: 4.3% vs. relevant-change: 5.7%), $t(18) = 2.84$, $p = .01$, $d = 0.48$. In the

DOI 10.1027/1618-3169/10.1027/1618-3169/a000511

difficult search condition, response-compatibility effects were only found for valid matching cues (relevant-repetition: 2.4% vs. relevant-change: 4.7%), $t(18) = 2.79$, $p = .01$, $d = 0.75$.

For the second five-way interaction, again, we first looked at the data from the lab-based version of the experiment. Under easy search and difficult conditions, we only found response-compatibility effects in valid trials, when the irrelevant feature changed; easy search: relevant-repetition: 1.9% vs. relevant-change: 3.3%, $t(12) = 2.73$, $p = .02$, $d = 0.79$; difficult search relevant-repetition: 4.1% vs. relevant-change: 6.3%, $t(12) = 2.74$, $p = .02$, $d = 0.61$. In the online version, results were slightly different: Under easy search conditions, response-compatibility effects were found for valid cues when the irrelevant feature changed (relevant-repetition: 2.3% vs. relevant-change: 4.4%), $t(18) = 2.82$, $p = .01$, $d = 0.68$. However, in addition, we also found response-compatibility effects for invalid cues when the irrelevant feature repeated (relevant-repetition: 3.6% vs. relevant-change: 4.3%), $t(18) = 2.3$, $p = .03$, $d = 0.33$. Under difficult search conditions, response-compatibility effects were only found for valid cues when the irrelevant feature repeated (relevant-repetition: 3.2% vs. relevant-change: 4.9%), $t(18) = 2.65$, $p = .02$, $d = 0.62$.

Compatibility between Cued Distractor and Target in Invalid Conditions

Response Times. We found a main effect of experiment version (44 ms difference; online version: 647 ms vs. lab-based version: 691 ms), $F(1, 30) = 4.88$, $p = .035$, $\eta_p^2 = .14$. This main effect was further qualified by its interaction with search difficulty, $F(1, 30) = 13.42$, $p = .001$, $\eta_p^2 = .31$. RTs in the difficult search condition were significantly faster in the online version (690 ms) than in the lab-based version (776 ms), 85 ms, $t(30) = 3.19$, $p < .01$, $d = 1.13$. No such difference was found in the easy search condition ($p = .7$).

DOI 10.1027/1618-3169/10.1027/1618-3169/a000511

Error Rates. Experiment version entered a two-way interaction with search difficulty also in error rates, $F(1, 30) = 14.77$, $p = .001$, $\eta_p^2 = .33$, a three-way interaction with cue match and irrelevant similarity, $F(1, 30) = 6.35$, $p = .017$, $\eta_p^2 = .17$, a four-way interaction with cue match, relevant similarity, and search difficulty, $F(1, 30) = 4.49$, $p = .042$, $\eta_p^2 = .13$, and, finally, a five-way interaction with search difficulty, cue match, relevant similarity, and irrelevant similarity, $F(1, 30) = 6.00$, $p = .02$, $\eta_p^2 = .17$. Since all factors interacted with each other, we only report the post-hoc analyses of this five-way interaction here: Taking a closer look at the lab-based version, we only found response-compatibility effects under the following easy search conditions: with matching cues both when the irrelevant feature was similar (relevant-dissimilar: 6.2% vs. relevant-similar: 1.7%), $t(12) = 3.74$, $p < .01$, $d = 1.27$, and when it was dissimilar (relevant-dissimilar: 4.8% vs. relevant-similar: 2.1%), $t(12) = 3.64$, $p < .01$, $d = 0.98$. In the online version, we found response-compatibility effects under easy search conditions with matching cues, but only when the irrelevant feature was similar (relevant-dissimilar: 6.2% vs. relevant-similar: 2.9%), $t(18) = 3.97$, $p < .001$, $d = 0.92$, and for non-matching cues when the irrelevant feature was similar (relevant-dissimilar: 3.4% vs. relevant-similar: 3.3%), $t(18) = 2.71$, $p = .014$, $d = 0.86$. In the difficult search condition of the online version of the experiment, response-compatibility were only found for matching cues when the irrelevant feature was similar (relevant-dissimilar: 4.2% vs. relevant-similar: 2.8%), $t(18) = 2.27$, $p = .036$, $d = 0.49$.

DOI 10.1027/1618-3169/10.1027/1618-3169/a000511