RUNNING HEAD: SHARED FEATURE ACCOUNT

Title Here

Sean Hughes, Simone Mattevelli, & Jan De Houwer

*Ghent University*

**Authors Note**

SH, SM, and JDH, Department of Experimental Clinical and Health Psychology, Ghent University. This research was conducted with the support of Grant BOF16/MET\_V/002 to JDH. Correspondence concerning this article should be sent to [sean.hughes@ugent.be](mailto:sean.hughes@ugent.be).

**Abstract**

*Keywords*:

**Title Here**

* Humans are exposed to a learning history that transforms them into ‘symbolic beings’ (e.g., Hughes, De Houwer, & Barnes-Holmes, 2016).
* For these organisms any cue in the environment, proximal or distal, can serve as a ‘symbol’ or be imbued with symbolic meaning (e.g., De Houwer & Hughes, 2016). Thus the topographical barrier between words and regularities melts away. Pairings, actions, or frequency can all function as symbols or contextual cues.
* Any common feature shared by stimuli can function as a contextual relational cue (symbol) indicating that those stimuli are related to one another, and as a result, a transfer of valence may take place from one stimulus to another.
* In most EC studies the shared feature is *contiguity*: the CS and US are similar with regard to their spatio-temporal properties. However – *in principle* – other common features may be enough for people to treat the stimuli as equivalent.
* We will explore this idea using *color* as a common feature. We created an EC procedure that contains a single CS and two USs – one positive and one negative. The CSs and USs are either presented in the same or different colors.
* We assume that CSs and USs that share a common color will produce larger EC effects than those that are presented in different colors (because people treat similar color as a cue indicating that a CS and US share other properties such as valence).

**Design of the Experiment**

**Between-subjects design:**

**Design**: 2 (*CS-US color matching:* CS1[CS2] matching positive[negative]USs vs. CS2[CS1] matching positive[negative]USs) between subject design. Method factors varied between participants:

* + - *Stimulus assignment*: CS1/CS2 identity assigned to same color as positive/negative words
    - *US identity*: set1 vs. set2 of positive and negative USs presented in CS1[CS2] trials.
    - *Order of measures*: IAT first vs. after self-reports
    - *IAT block order* (consistent vs. inconsistent with learning phase).
* EC 🡪 Evaluative measures 🡪 Exploratory Questions

- Stimulus:

Two nonsense words (MORAG and STRUAN) will serve as CS1 and CS2.

Six positive (Rainbow, pleasure, smile, love, paradise, joy) and six negative adjectives (war, cancer, hate, hell, misery, vomit), divided in four sets (two sets of positive words and two sets of negative words) will serve as USs.  
**Participants:**

Data-collection will be via the Prolific website (prolific.achttps://prolific.ac/). We will stop data-collection as soon as 110 participants have completed the experiment on the Prolific website. This will allow us to have good power (> 0.80) to observe an EC effect (*d* = 0.50) driven by US-CS color matching, at alpha = 0.05.

**Procedure:**

**EC training phase**: Participants receive three blocks of 16 trials (48 total) consisting of two different types of trials: one type of trials wherein CS1 is presented in the same color as positive words, and another trial in which CS2 is presented in the same color as negative words. Note that each trial will contain three stimuli simultaneously presented onscreen: a neutral word (MORAG or STRUAN) and a positive and negatively valenced adjective. All three stimuli will initially be presented in white. Then after 3000ms the CS will change to one color (e.g., blue). Depending on the CS present on that trial, one US will also turn that same color (e.g., blue) whereas the other will change to a different color (e.g., purple). The stimuli will remain onscreen for another 3000ms before all stimuli being removed, an inter-trial interval of 1250ms and the next trial. Stimuli color will varied across each trial, so that none of the colors can assume any specific positive or negative value. Four different colors (i.e., blue, green, yellow and purple) will be used.

*Same Color (CS1 trial) Same Color (CS2 trial)*

**JOY HELL**

**MORAG STRUAN**

**SMILE**

**CANCER**

*IAT.* Participants perform an IAT measuring implicit evaluations of CS1 vs. CS2:   
The IATs involve categories “CS1”[“CS3”] and “CS2” [“CS4”] and “Good” and “Bad”  
  
IAT Procedure:

a. Instructions: “In In the next part you will have to categorize items into groups as fast as you can.”

b. 20 practice trials sorting CS1 and positive words using one key.   
  
c. 20 practice trials sorting CS2 and positive words using one key.

d. 20 test trials CS1 and positive words using one key.

e. 40 test trials CS2 and positive words using one key.

f. 20 test trials CS1 and positive words using one key.

g. 40 test trials CS2 and positive words using one key.

NOTE: IAT stimuli are:

|  |  |
| --- | --- |
| CS1 |  |
| CS2 |  |
| Positive | /1 = "Fantastic"  /2 = "Great"  /3 = "Nice"  /4 = "Good"  /5 = "Pleasant"  /6 = "Wonderful"  /7 = "Amazing"  /8 = "Happy" |
| Negative | /1 = "Terrible"  /2 = "Disgusting"  /3 = "Nasty"  /4 = "Horrible"  /5 = "Sick"  /6 = "Awful"  /7 = "Sad"  /8 = "Unpleasant" |

*Explicit Attitude:* Participants give explicit ratings for the two CSs by answering the question:

*“Please rate the above item using the scale below*  
options: -5 =Negative, 5= Neutral, +5= Positive

options: -5 =I Dislike it, 5= Neutral, +5= I Like it

options: -5 =Bad, 5= Neutral, +5= Good

options: -5 =Unpleasant, 5= Neutral, +5= Pleasant

*Intention Ratings:* Participants are presented with two brand products labeled with either CS1 or CS2. They are asked to indicate which of these products they would try and given the following options: “I would try CS1, I would try CS2, I would try CS1 and CS2, I would try neither, I don’t know”.

Finally, participants answer the following questions about the EC task:

*Contiguity memory:* CS1, CS2, was presented onscreen and participants asked: “In the first part of the experiment (when words appeared initially in white and then switched their color) MORAG/STRUAN was always presented with two words. What was the meaning of those words?” ("Both words always had a positive meaning", " Both words always had a negative meaning", " One words always had a positive meaning and the other one a negative meaning" "I don’t remember")

*Color memory:* CS1, CS2, was presented onscreen and participants asked: “In the first part of the experiment (when words appeared initially in white and then switched their color) did MORAG/STRUAN switch to the same color as” ("The positive word that was also on the screen", "The negative word that was also on the screen", "I don’t remember")

*Manipulation check* to ensure that participants did not write down the contingencies during the learning phase: “Think back to the first part of the experiment (i.e., when three words were paired onscreen). Did you ever take notes (or write down) what happened in order to help you figure out what was going on? Please be honest here (you will receive payment regardless of what you say).”

*Demand Compliance (explicit):* Earlier you rated MORAG and STRUAN as being either positive, neutral, or negative. Did you base your ratings NOT on how you actually felt about those words but ONLY on what you thought the researchers wanted you to say?" ("Yes", "No", "I don’t know")

*Demand compliance (implicit).* Earlier you completed the Implicit Association Test (see below). Did you base your performance in that task NOT on your best efforts to perform the categorizations as quickly and accurately as possible but on your attempt to influence your speed or accuracy in order to go along with what you thought the researchers wanted you to feel about the words? ("Yes", "No", "I don’t know")

*Reactance. (Explicit).* Earlier you rated MORAG and STRUAN as being either positive, neutral, or negative. Did you consciously resist what you thought the researchers wanted you to feel about those words?" ("Yes", "No", "I don’t know")

*Reactance (implicit).* Earlier you completed the Implicit Association Test (see below). Did you try to influence your speed or accuracy in order to consciously resist what you thought the researchers wanted you to feel about those words" ("Yes", "No", "I don’t know")

*Hypothesis awareness*. During the first part of the study, did you notice that the color of MORAG and STRUAN switched to the same color as either positive or negative words? Please be honest here"

*Influence awareness*. Did this influence how you responded to question about your liking of MORAG and STRUAN? Please be honest here"

**Data preparation**

107 participants (59 females, *Mage =* 33.39*, SD =* 8.47) took part to the study on Prolific. We excluded data from one participants who did not complete the entire session. Moreover the data of participants who had IAT error rates for any of the IATs above 30% across the entire task, or above 40% for any one of the four critical blocks or for participants who complete more than 10% of IAT trials faster than 400 ms (*N* = 3) were excluded from the analyses. This led to a final sample of 103 participants.

**Results**

We first looked at the effect of CS-US color matching on implicit attitude change. We ran a one-way ANOVA and found a significant effect of color matching on the IAT score, *F*(1, 101) = 43.91, *p* < .001, η2partial **=** .30. Participants that saw CS1 appearing in the same color as positive words showed higher IAT score (revealing a preference for CS1 over CS2). We conducted the same analysis using the differential explicit score as dependent variable. As we did for the IAT score, also for explicit attitude higher scores indicated and explicit preference for CS1 over CS2. We found a main effect of CS-US color matching on explicit attitudes, *F*(1, 101) = 70.57, *p* < .001, η2partial **=** .41. Again, participants expressed higher preference for the CS that was presented in the same color as positive USs. Then, we checked whether the color matching manipulation also resulted in a significant effect on intentions towards the two CSs. We looked at the difference in the proportion of participants who intended to purchase CS1 and CS2 in the two conditions. The proportion of responses in favor of CS1 was higher when CS1 matched the color of the positive USs (.40) than when it matched the negative ones (.14, *z* = 3.00, *p* = 0.002). The opposite pattern emerged for the proportion of choices in favor of CS2 (.09 vs. .43, *z* = -3.90, *p* < 0.001).

*Color Matching memory*. A color memory score that ranged from 0 to 2 was calculated. A 0 score indicated that participants’ response was incorrect to both CS1 and CS2 color memory question (*N* = 23), 1 indicated at least a correct response (*N* = 6) and 2 indicated that both the questions were answered correctly (*N* = 74). We conducted a moderation analysis to see whether color memory score qualified the impact of our manipulation on implicit and explicit attitude change. We found that color memory moderated the impact of the manipulation on both implicit and explicit attitude change, *b* = .27, *p* < .001 and *b* = .34, *p* < .001. In both cases, the effect of the manipulation increased as a function of correct color memory.

*Contiguity awareness*. We found that 66% of participants (*N* = 68) responded correctly to the contiguity questions about CS1 and CS2 (i.e., “One words always had a positive meaning and the other one a negative meaning”). Contiguity responses in which at least one of the two answers (CS1- and CS2-USs contiguity) indicated the opposite contiguity (e.g., “Both words always had a negative meaning” when the CS was matched with positive stimuli), and when participants did not remember at least one of the two (*N* = 11). We noticed however that another type of incorrect contiguity response emerged for a considerable proportion of participants (23%): 18 participants remembered that both the CSs were paired with 2 US stimuli of the same valence (i.e., “Both words always had a negative meaning” for CS matched in color with negative stimuli; “Both words always had a positive meaning” for CS matched in color with positive stimuli), and 5 participants did so for at least one of the CS while they responded correct to the other one. Comparing the effect of the manipulation across participants with either correct contiguity or color-matching-driven contiguity memory resulted in a significant interaction only on explicit attitude change, *F*(1, 96) = 5.36, *p* = .006, η2partial **=** .10, such that the effect was stronger for participants with color-matching-driven contingency. Therefore, participants memory concerning the valence of the two USs presented on screen was affected by the valence of the US that matched the color of the CS at issue. This might imply that, at least for some participants, color matching affect evaluative scores by influencing how stimuli-pairings are processed by the individuals.

*Hypothesis (color) and influence Awareness.* We looked at participants’ response to the color awareness question (i.e., During the first part of the study, did you notice that the color of MORAG and STRUAN switched to the same color as either positive or negative words?). We found that 26% of the participants did not notice this. We then re-ran the analyses considering hypothesis awareness as an additional factor in a 2x2 ANOVA. We found no significant interaction on implicit, *F*(1, 98) = 1.29, *p* = .260, and a significant one on explicit, *F*(1, 98) = 9.48, *p* = .003, η2partial **=** .09. Among participants who did actually notice the switching in color, we found that 65% of them reported that the color switching influenced the way the evaluated the CSs, while for the remaining 35% it did not. Results confirmed that on those who explicitly indicated they have been influenced by the color switching (*N =* 49), the impact of our manipulation was stronger on both implicit, *F*(1, 47) = 40.46, *p* < .001, η2partial **=** .46, and explicit attitude change, *F*(1, 47) = 189.80, *p* < .001, η2partial **=** .80.

*Reactance.* We had 17% of reactant participants for the explicit measures and 20% for the IAT. The exclusion of reactant participants did not affect the magnitude of the effect on either implicit or explicit attitude change.

**Conclusion**

We obtained evidence to support the shared feature account. A neutral CS was presented presented with both positive and negative USs. Whenever a CS and US shared the same color a corresponding change in the CSs valence (in line with the US) emerged, such that CSs that appeared in the same color as positive USs became positive, while CSs that appeared in the same color as negative USs became negative. We obtained evidence for this on measures of implicit and explicit evaluations as well as behavioral intention questions.

**Experiment 2**

* Experiment 1 confirmed our hypothesis, showing that CSs acquire the valence of the US that switches to the same color as the CS. This effect was evident on implicit and explicit measures of evaluation and behavioral intentions.
* Upon reflection, however, the fact that the CS acquired the valence of the US that matched it in color might be due to (a) the regularity in the *mere* *presence* of two stimuli of the same color, or (b) the regularity in the *simultaneous* *switching* of two stimuli to the same color.
  + Specifically, the EC effect could be moderated by the fact that two (of the three) stimuli shared the same color (as we suspect). But it is also possible that participants’ attention was captured by the *change in color* and that they did not attend to the third stimulus onscreen (i.e., they only perceived, attended to, or remembered the presence of two stimuli and not three). If this was the case, then the EC was less to do with a common feature and more to do with the fact that participants were only attending to two, rather than three, stimuli.
* We will explore this in Experiment 2. Rather than being initially presented in white, all the stimuli will appear in one color at the beginning of the trial and then, only the unmatched US will switch to a different color, while the CS and the matched US will maintain the original color.
* In this way, we can investigate whether eliminating simultaneous switching from color matching has an impact in the observed transfer of valence.

Two potential outcomes are plausible:

* *Color matching hypothesis.* We could expect a main effect of US-CS color matching on the subsequent liking towards the CSs, such that the CSs will acquire the valence of the US that maintains the same color as it.
* *Switching salience hypothesis.* People might perceive the switch in color as a way to highlight that one of the two USs presented on screen is actually connected to the CS. If this is the case, then the CS should acquire the valence carried by the highlighted US, leading to the opposite effect as we originally predicted.
* Note: in both cases the EC effect would still be inferential in nature. But the nature of the inferences made about color and the CS-US relationship imply different outcomes.

**Design**

**Between-subjects design:**

**Design**: 2 (*CS-US color matching:* CS1[CS2] matching positive[negative]USs vs. CS2[CS1] matching positive[negative]USs) between-subject design. Method factors varied between participants:

* + - *Stimulus assignment*: CS1/CS2 identity assigned to same color as positive/negative words
    - *US identity*: set 1 vs. set 2 of positive and negative USs presented in CS1[CS2] trials.
    - *Order of evaluative measures*: IAT before vs. after self-reports
    - *IAT block order* (consistent vs. inconsistent with learning phase).

**Stimuli.** Two nonsense words (MORAG and STRUAN) will serve as CS1 and CS2. Six positive (*Rainbow, Pleasure, Smile, Love, Paradise, Joy*) and six negative adjectives (*War, Cancer, Hate, Hell, Misery, Vomit*), divided in four sets (two sets of positive words and two sets of negative words) will serve as USs.  
**Participants.** Data-collection will be via the Prolific website (prolific.achttps://prolific.ac/). We will stop data-collection as soon as 110 participants have completed the experiment on the Prolific website. This will allow us to have good power (> 0.80) to observe an EC effect (*d* = 0.50) driven by US-CS color matching, at alpha = 0.05.

**Procedure.**  EC 🡪 Evaluative measures 🡪 Exploratory Questions

*EC training phase*. Participants receive three blocks of 16 trials (48 total) consisting of two different types of trials: one type of trial wherein CS1 is presented in the same color as positive words, and another trial in which CS2 is presented in the same color as negative words. Note that each trial will contain three stimuli simultaneously presented onscreen: a neutral word (MORAG or STRUAN) and positive and negatively valenced words. All three stimuli will initially be presented in the same color (e.g., blue). Then after 3000ms, depending on the CS present on that trial, one US will turn change to a different color (e.g., purple), while the CS and the other US will maintain the same color (e.g., blue). The stimuli will remain onscreen for another 3000ms before all stimuli are removed, an inter-trial interval of 1250ms, and the next trial. Stimulus color will be varied across each trial, so that none of the colors can assume any specific positive or negative value. Four different colors (i.e., blue, green, yellow and purple) will be used.

*Same Color (CS1 trial) Same Color (CS2 trial)*

**JOY HELL**

**MORAG STRUAN**

**SMILE**

**CANCER**

*Implicit attitude*

Participants will perform an IAT measuring implicit evaluations of CS1 vs. CS2:   
The IATs involve categories “CS1” and “CS2” and “Good” and “Bad”  
  
IAT Procedure:

a. Instructions: “In the next part you will have to categorize items into groups as fast as you can.”

b. 20 practice trials sorting CS1 on the left and CS2 on the right.   
  
c. 20 practice trials sorting positive words on the left and negative words on the right.

d. 20 test trials CS1 and positive words using one key and CS2 and negative words using another key.

e. 40 test trials CS1 and positive words using one key and CS2 and negative words using another key.

f. 20 practice trials CS2 on the left and CS1 on the right.

g. 20 test trials CS2 and positive on the left and CS1 and negative on the right.

h. 40 test trials CS2 and positive on the left and CS1 and negative on the right.

NOTE: IAT stimuli are:

* + - CS1
    - CS2
    - Positive words: *Fantastic, Great, Nice, Good, Pleasant, Wonderful, Amazing, Happy*
    - Negative words: *Terrible, Disgusting, Nasty, Horrible, Sick, Awful, Sad, Unpleasant*

*Explicit attitude*. Participants give explicit ratings of the two CSs by answering the question:

*“Please rate the above item using the scale below*  
options: -5 =Negative, 5= Neutral, +5= Positive

options: -5 =I Dislike it, 5= Neutral, +5= I Like it

options: -5 =Bad, 5= Neutral, +5= Good

options: -5 =Unpleasant, 5= Neutral, +5= Pleasant

*Intention measure*. Participants are presented with two brand products labeled with either CS1 or CS2. They are asked to indicate which of these products they would try and given the following options: “I would try CS1, I would try CS2, I would try CS1 and CS2, I would try neither, I don’t know”.

Finally, participants answer the following exploratory questions about the EC task:  
  
 *CS-USs Contiguity memory*: CS1, CS2, was presented onscreen and participants asked: “In the first part of the experiment was MORAG[STRUAN] presented together with.”

"positive words only", "negative words only", "both positive and negative words", "neither positive or negative words" "I don’t remember")

*Color switch awareness*. “During the first part of the study, with the colored words, did you notice that the color of one of the two words presented on the right side of the screen switched, while the word on the left side of the screen (MORAG/STRUAN) stayed the same? Please be honest here "

*Color switch influence*. “Did this influence how you responded to MORAG and STRUAN? Please be honest here"

*Color switch contingencies*: “In the first part of the experiment, when MORAG[STRUAN] appeared on the screen, which words switched to a different color”

("The positive words", "The negative words", "I don’t remember")

*Manipulation check*. “Think back to the first part of the experiment (i.e., when three words were paired onscreen). Did you ever take notes (or write down) what happened in order to help you figure out what was going on? Please be honest here (you will receive payment regardless of what you say).”

*Demand Compliance (explicit):* Earlier you rated MORAG and STRUAN as being either positive, neutral, or negative. Did you base your ratings NOT on how you actually felt about those words but ONLY on what you thought the researchers wanted you to say?"

("Yes", "No", "I don’t know")

*Demand compliance (implicit).* Earlier you completed the Implicit Association Test (see below). Did you base your performance in that task NOT on your best efforts to perform the categorizations as quickly and accurately as possible but on your attempt to influence your speed or accuracy in order to go along with what you thought the researchers wanted you to feel about the words?

("Yes", "No", "I don’t know")

*Reactance. (Explicit).* Earlier you rated MORAG and STRUAN as being either positive, neutral, or negative. Did you consciously resist what you thought the researchers wanted you to feel about those words?"

("Yes", "No", "I don’t know")

*Reactance (implicit).* Earlier you completed the Implicit Association Test (see below). Did you try to influence your speed or accuracy in order to consciously resist what you thought the researchers wanted you to feel about those words"

("Yes", "No", "I don’t know")

**Data preparation**

118 participants (67 females, *Mage =* 32.3*, SD =* 8.6) took part to the study on Prolific. We excluded data from nine participants who did not complete the entire session. The data of participants who had IAT error rates above 30% across the entire task, or above 40% for any one of the four critical blocks, or for participants who responded faster than 400ms on more than 10% of IAT trials (*N* = 3) were also excluded from the analyses. This led to a final sample of 106 participants.

**Descriptive Statistics**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Stimulus** | **CS1** | | **CS2** | | **Total** | |
|  | **Yes** | **No** | **Yes** | **No** | **Yes** | **No** |
| CS-US Contingency | 67% | 33% | 70% | 30% |  |  |
| Aware of US Color Switch Contingency | 66% | 34% | 68% | 32% |  |  |
| Color Change Awareness |  |  |  |  | 87% | 13% |
| Color Change Influence |  |  |  |  | 49% | 51% |
| Demand Explicits |  |  |  |  | 12% | 88% |
| Demand Implicits |  |  |  |  | 14% | 86% |
| Reactance Explicits |  |  |  |  | 16% | 84% |
| Reactance Implicits |  |  |  |  | 9% | 91% |
| Wrote Down the Contingencies |  |  |  |  | 10% | 90% |

**Results**

**Self-reports.** We first looked at the effect of CS-US color matching on explicit attitude change. We calculated a difference score by subtracting explicit ratings for the stimulus that stayed the same color as negative words (CS2) from the stimulus that stayed the same color as positive words (CS1). Positive scores indicate a preference for CS1 over CS2 (i.e., for the stimulus that remained the same color as a US) whereas negative scores indicate a preference for CS2 over CS1 (i.e., for a relationship between the CS and the US that changed its color). Or put another way, a positive EC effect would support the common feature account whereas negative scores would support a salience account. A one sample t-test revealed a significant EC effect supporting the salience over the common feature account (*M* = -1.9, *SD* = 5.8), *t*(105) = 3.48, *p* < .001, *d* = 0.34. That is, participants preferred CS2 over CS1, suggesting that they liked the previously stimulus that (a) remained the same color as negative USs, but where (b) the positive stimulus changed color more than the stimulus that (c) remained the same color as positive USs, but where (d) the negative stimulus changed color (i.e., their attention was captured more by the changing color more than the commonality in color between stimuli).

**IAT**. We conducted the same analysis using the IAT score as dependent variable. Once again, a positive IAT score indicates a preference for CS1 over CS2 where as a negative score indicates the opposite. Once again we found a significant EC effect for CS2 over CS1, supporting the salience over the common feature account (*M* = -0.18, *SD* = 0.57), *t*(105) = 3.35, *p* < .001, *d* = 0.32.

*Color Switch contingency memory*. We assessed if participants were aware of the valence of the US that switched color during each trial. We calculated a color switch contingency memory score, ranging from 0 to 2, based on responses to the following question: “In the first part of the experiment, when CS1[CS2] appeared on the screen, which words switched to a different color”. A 0 score indicated that people responded incorrectly to both CS1 and CS2 (*N* = 35), 1 indicated at least one correct response (*N* = 6) and 2 indicated that both the questions were answered correctly (*N* = 65). If we code individuals as having passed or failed that test, and add this factor to a one-way ANOVA, we see a significant effect for explicit, *F*(1, 105) = 9.00, *p* = .003, and implicit evaluations, *F*(1, 105) = 14.37, *p* < .001, such that both evaluations increase in magnitude, in a negative direction when people can correctly indicate the valence of the US that changed color during a given trial.

*CS-US* *contiguity awareness*. We found that 65% of participants (*N* = 69) responded correctly to the questions about the CS-US contingencies (i.e., they selected “both positive and negative words” when presented with the following question: “In the first part of the experiment was CS1[CS2] presented together with… positive words only, negative words only, both positive and negative words, neither positive or negative words, or I don’t remember”). Of the remaining 35% (n = 37), seven (7%) responded to one of the CSs correctly and the other incorrectly. The other 30 participants responded incorrectly on both questions, either indicating that they could not remember (n = 11) or answering both questions incorrectly. Coding participants as either having passed or failed the memory test, and adding this as a factor in a one-way ANOVA, revealed a descriptively larger effect for explicit, *F*(1, 105) = 3.02, *p* = .08, but not implicit scores, *F*(1, 105) = 0.49, *p* = .48, when participants were aware of the contingency between CSs and USs.

*Hypothesis (color) and influence Awareness.* We looked at participants’ response to the color awareness question (i.e., *During the first part of the study, with the colored words, did you notice that the color of one of the two words presented on the right side of the screen switched, while the word on the left side of the screen (MORAG/STRUAN) stayed the same?*). We found that 87% of participants did notice this whereas 13% did not. We re-analyzed the data considering hypothesis awareness as a factor in a one-way ANOVA. We found no interaction on either implicit, *F*(1, 105) = 0.17, *p* = .68, or explicit evaluations, *F*(1, 105) = 2.45, *p* = .12. Among participants who noticed the color switch, we found that only 54% of them reported that the color switching influenced the way they evaluated the CSs, whereas the remaining 46% said it did not. For those who explicitly reported being influenced by the color switching (*N =* 50), the impact of our manipulation was descriptively stronger on both implicit, *F*(1, 91) = 2.03, *p* = .16, and explicit attitude change, *F*(1, 91) = 2.96, *p* = .09, again in the negative direction (i.e., a preference for CS2 over CS1).

*Demand*. We had 12% of demand compliant participants for the explicit measures and 14% for the IAT. The exclusion of these participants did not affect the magnitude of implicit or explicit attitude change. If anything both effects become stronger in the negative direction.

*Reactance.* We had 16% of reactant participants for the explicit measures and 9% for the IAT. The exclusion of reactant participants did not affect the magnitude of implicit or explicit attitude change. If anything both effects become stronger in the negative direction.

**Conclusion**

In Experiment 1 we started with a situation where all three stimuli (CS1, USpos, and USneg) were in white and then CS1 and USpos would change. So we first highlighted that all three stimuli were the same and then only later did we highlight that CS1 and USpos shared a common color.

In Experiment 2 we start with a situation were all three stimuli (CS1, USpos, and USneg) already shared a color and then one of the USs changes into a different color. From the data it seems that participants tend to see the US that changes color as being more related to the CS than the US that stays in the same color. This is true for both explicit and implicit evaluations. When I look at the data I see several patterns. First there seems to be two distinct groups of participants in our sample. Those that form the inference “the CS and US share a color 🡪 they are related” and a second group who say “they have just changed the color of one of the USs…they must be pointing out that it is important or related to the CS in some way”. This latter group exerts more of an impact on responding than the first. Participants reported something along these lines during the color switch awareness question. Also you can see it in the following graphs (i.e., explicit evaluations on Y axis and participant number on x axis). For instance:

When MORAG Stays same color as USpos while USneg changes color then half of the participants say that Morag is good and the other half say it is bad.

When STRUAN Stays same color as USneg while USpos changes color most say it is good whereas some say it was bad

When Struan stays the same color as USpos while USneg changes color then half say it is good and the other half say it is bad

Finally, when Morag stays the same color as USneg while USpos changes color most say it is good and some say it is bad.

So it seems that the training procedure allows for two different types of evaluative inferences to be made on the basis of color change. I think this happened for several reasons.

* First, in *Experiment 1*, our instructions were clear (“In the next part of the study…you will encounter two new words: MORAG and STRUAN. These words will appear on the screen together with two other words. The new word (MORAG or STRUAN) and the other words will initially appear in white. Then the color of the three words will change”). In *Experiment 2* we said the following (“In the first part you will see two new words: MORAG and STRUAN. These words will appear onscreen together with two other words. The new word (MORAG or STRUAN) and other words will initially appear in one color. **Then the color of one of the words will change**…**Please pay close attention to the color of each word and how they change**”).

In short, Experiment 1 instructions and the task make it easy to form the inference that shared color is important. Experiment 2 instructions and the task highlight that people should pay attention to the *change in color – not the fact that stimuli remained in the same color*. This may have caused people to shift their attention to the changing color and consider it more diagnostic about CS meaning than the same colored US.

* So how can we get people to form the inference we want them to? I think that replicating the experiment with new instructions that focus people’s attention on the fact that certain stimuli remain in the same color might produce the expected effects.

**Experiment 3**

* In Experiment 3 we will attempt to replicate and extend Experiment 2. Specifically, all aspects of the study will remain the same. The only difference will be the task instructions which will now emphasize to participants that the CS and US will remain in the same color (rather than telling them that the CS and US will change color). We expect that these new instructions may cause most people to form the inference that the CS and US are related (because they share a common feature – color) and thus lead to assimilative EC effects.

**Design**

**Participants.** Data-collection will be via the Prolific website (prolific.achttps://prolific.ac/). We will stop data-collection as soon as 110 participants have completed the experiment on the Prolific website. This will allow us to have good power (> 0.80) to observe an EC effect (*d* = 0.50) driven by US-CS color matching, at alpha = 0.05.

Procedure. Identical to Study 2 with the exception of the instructions used.

“Part I: You are going to see a new word appear on the screen (i.e., MORAG or STRUAN). MORAG or STRUAN will appear on the left of the screen. Two other words will appear on the right. MORAG or STRUAN and other words will first appear in the same color. MORAG or STRUAN will STAY THE SAME COLOR as one of the words on the right. \*\*Please pay close attention to the colors of the words. You will be asked some questions about this later on.\*\*” After each block they were told that “^^ Ok. We will do this task again. Please press the space bar to continue. REMEMBER: MORAG or STRUAN will stay the same color as one of the words on the right.”

**Results**

**Data preparation**

118 participants (70 females, *Mage =* 28.19*, SD =* 6.08) took part to the study on Prolific. We excluded data from 9 participants who did not complete the entire session. The data of participants who had IAT error rates above 30% across the entire task, or above 40% for any one of the four critical blocks, or for participants who responded faster than 400ms on more than 10% of IAT trials (*N* = 12) were also excluded from the analyses. This led to a final sample of 97 participants.

**Descriptive Statistics**

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
| **Stimulus** | **CS1** | | **CS2** | | **Total** | |
|  | **Yes** | **No** | **Yes** | **No** | **Yes** | **No** |
| CS-US Contingency | 74% | 26% | 72% | 28% |  |  |
| Aware of US Color Switch Contingency | 69% | 31% | 68% | 32% |  |  |
| Color Change Awareness |  |  |  |  | 91% | 9% |
| Color Change Influence |  |  |  |  | 42% | 58% |
| Demand Explicits |  |  |  |  | 23% | 77% |
| Demand Implicits |  |  |  |  | 19% | 81% |
| Reactance Explicits |  |  |  |  | 14% | 86% |
| Reactance Implicits |  |  |  |  | 11% | 89% |
| Wrote Down the Contingencies |  |  |  |  | 8% | 92% |

**Results**

**Self-reports.** We first looked at the effect of CS-US color matching on explicit attitude change. We calculated a difference score by subtracting explicit ratings for the stimulus that stayed the same color as negative words (CS2) from the stimulus that stayed the same color as positive words (CS1). Positive scores indicate a preference for CS1 over CS2 (i.e., for the stimulus that remained the same color as a US) whereas negative scores indicate a preference for CS2 over CS1 (i.e., for a relationship between the CS and the US that changed color). Or put another way, a positive EC effect would support the common feature account whereas negative scores would support a salience account. A one sample t-test revealed no significant EC effect (*M* = -0.17, *SD* = 5.9), *t*(96) = -.29, *p* = .77.

**IAT**. We conducted the same analysis using the IAT score as dependent variable. Once again, a positive IAT score indicates a preference for CS1 over CS2 whereas a negative score indicates the opposite. We found a significant EC effect supporting the common feature over the salience account (*M* = 0.18, *SD* = 0.53), *t*(96) = 3.25, *p* = 0.002, *d* = 0.33.

*Color Switch contingency memory*. 30 people responded incorrectly to both CS1 and CS2, 1 indicated at least one correct response, and 66 indicated that both the questions were answered correctly. When coding these individuals as having passed or failed that test, and adding this factor to a one-way ANOVA, we did not find an effect for implicit, *F*(96) = 0.01, *p* = .93, nor explicit evaluations, *F*(96) = 1.14, *p* = .29.

*CS-US* *contiguity awareness*. We found that 72% of participants (*N* = 70) responded correctly to the questions about the CS-US contingencies (i.e., they selected “both positive and negative words” when presented with the following question: “In the first part of the experiment was CS1[CS2] presented together with… positive words only, negative words only, both positive and negative words, neither positive or negative words, or I don’t remember”). Of the remaining 28% (n = 27), two (7%) responded to one of the CSs correctly and the other incorrectly. The other 25 participants responded incorrectly on both questions, either indicating that they could not remember (n = 5) or answering both questions incorrectly. Coding participants as either having passed or failed the memory test, and adding this as a factor in a one-way ANOVA, revealed no effect for explicit, *F*(96) = 0.63, *p* = .43, or implicit scores, *F*(96) = 1.32, *p* = .25.

*Hypothesis (color) and influence Awareness.* We looked at participants’ response to the color awareness question (i.e., *During the first part of the study, with the colored words, did you notice that the color of one of the two words presented on the right side of the screen switched, while the word on the left side of the screen (MORAG/STRUAN) stayed the same?*). We found that 91% of participants did notice this whereas 9% did not. We re-analyzed the data considering hypothesis awareness as a factor in a one-way ANOVA. We found no interaction on either implicit, *F*(96) = 0.04, *p* = .84, or explicit evaluations, *F*(96) = 0.96, *p* = .33. Among participants who noticed the color switch, we found that only 44% of them reported that the color switching influenced the way they evaluated the CSs, whereas the remaining 56% said it did not. 41 participants explicitly reported being influenced by the color switching. The impact of our manipulation was strong on implicit attitude change, *F*(96) = 6.32, *p* = .02, but not on explicit attitude change, *F*(96) = .27, *p* = .61, as participants who were aware about the influence showed a stronger implicit preference for CS1 over CS2.

*Demand*. We had 23% of demand compliant participants for the explicit measures and 19% for the IAT. The exclusion of these participants did not affect the magnitude of implicit or explicit attitude change.

*Reactance.* We had 14% of reactant participants for the explicit measures and 11% for the IAT. The exclusion of reactant participants did not affect the magnitude of implicit or explicit attitude change.

**Conclusion**