**General Information**

* 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 symbol indicating that those stimuli are equivalent/similar to one another, and as a result, a transfer of valence may take place from one stimulus to another.
* In most EC studies the common 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).
* 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 of Experiment 2**

**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 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 stimulus that remained the same color as negative USs, but where the positive stimulus changed color over the stimulus that remained the same color as positive USs, but where the negative stimulus changed color.

**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 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.

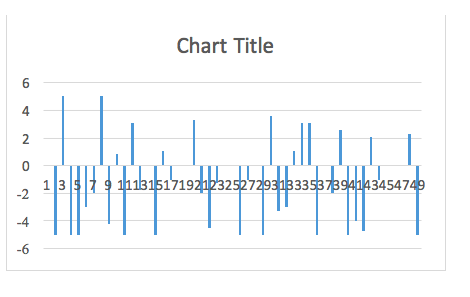
*Reactance.* We had 16% of reactant participants for the explicit measures and 9% for the IAT. The exclusion of reactant participants did not affect the magnitude of implicit or explicit attitude change. If anything both effects become stronger in the negative direction.

**Conclusion**

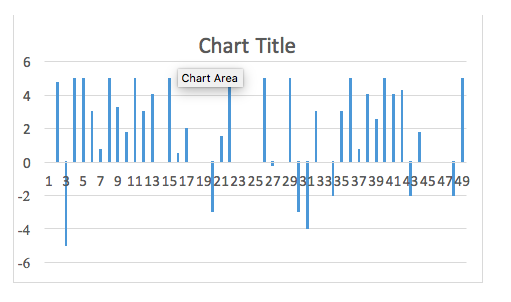
Just to briefly recap. In Experiment 1 we started with a situation where all three stimuli (CS1, USpos, and USneg) were in white and then two (CS1 and USpos) would change. So we first highlighted that all three stimuli were the same and then only later did we highlight that two of them shared a common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

In Experiment 2 we start with a different situation. Specifically, one where all three stimuli (CS1, USpos, and USneg) are already sharing a color and then one of them (US) changes to 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. 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 this US is important or related to the CS in some way”. This latter group exerts more of an impact on overall responding than the first. Participants reported something along these lines during the color switch awareness question (see analysis files). Also you can see it in the following graphs (i.e., explicit evaluations on Y axis and participant number on x axis). For instance:

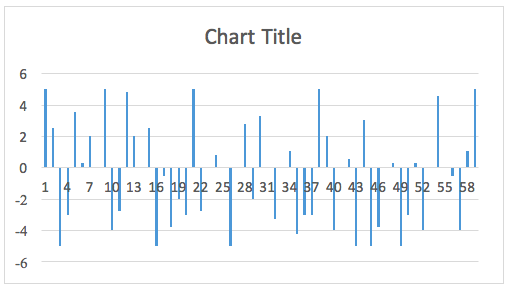
When MORAG stays in the same color as USpos (while USneg changes color) then half of the participants say that Morag is good and the other half say it is bad.



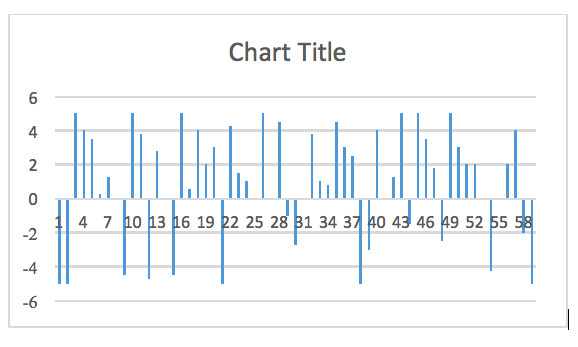
When STRUAN stays in the same color as USneg (while USpos changes color) most say it is good whereas some say it is bad.



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



Finally, when Morag stays in 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 is happening 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”). Yet 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’s instructions and the EC task itself make it relatively easy to form the inference that common color is important. Experiment 2’s instructions and the EC 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 a similar colored US.

* So how can we get people to form the inference we want them to? I think we can do so by replicating Experiment 2 with new instructions that focus people’s attention on the fact that certain stimuli remain in the same color could do so.
* However, we also had another way to test the alternative explanations of our initial (Experiment 1) findings. For instance, we could replicate Experiment 1’s design with Jan’s suggestions about a memory test during the EC phase (to show that people are actively processing all the stimuli that are onscreen). But it would be nice to show that we can also get the effect with this procedure as well…

For instance, what if instead of changing to a different color, the US always changed to white instead? Or simply went transparent. Basically, de-emphasizing the importance of the US that changes might work in Experiment 2.

* Note: regardless, it shows that the inferences that people make about common or different colors moderate the impact of the CS-US relationship on liking. So still consistent with a propositional perspective, I believe.

Minor points:

* People reported that three blocks of 16 trials was a lot. Could consider shortening.
* Think about a change in CS identity as scores were generally smaller for Struan than Morag