Acta Psychologica 203 (2020) 103011

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

# Learning the affective value of people: More than affect-based mechanisms

Chiara Ferrari[a,*], DongWon Oh[b], Brandon P. Labbree[b], Alexander Todorov[b]

[a] Department of Brain and Behavioral Sciences, University of Pavia, Pavia, Italy
[b] Department of Psychology, Princeton University, Princeton, USA

ARTICLE INFO

*Keywords:*
Affective learning
Valence
Moral values
Person attribution

ABSTRACT

People's ability to learn about the affective value of others is impressive. However, it is unclear whether this learning solely reflects general affect-based processes or a mixture of affect-based and person-attribution processes. Consistent with the former possibility, people's ability to learn the affective value of people and places have been shown to be comparable (Falvello, Vinson, Ferrari, & Todorov, 2015). To investigate whether general affect-based processes are sufficient to account for this kind of learning, we presented participants with images paired with valenced statements that were either relevant (e.g., a person statement with a person image) or irrelevant (e.g., a person statement with a non-person image). After this presentation, participants evaluated the goodness or badness of the images. In Experiment 1, we found that the learning effects for faces and places were comparable and occurred only when the statements were relevant. However, when we presented the images with multiple statements of the same valence (Experiments 2–4), we found that places acquired affective value from both relevant and irrelevant statements. In contrast, faces were less likely to acquire affective value from irrelevant statements. Our findings suggest that although general affect-based processes might be sufficient to account for affective learning of places, affective learning of faces might involve both affect-based and person-attribution processes.

## 1. Introduction

People are fundamentally motivated to evaluate other people (Uhlmann, Pizarro, & Diermeier, 2015) and show an impressive ability to do it. This is especially true in the moral domain, which is the social dimension dominating person evaluation (e.g., Landy, Piazza, & Goodwin, 2018; Uhlmann et al., 2015; Wojciszke, Bazinska, & Jaworski, 1998). People make judgments about others' trustworthiness from faces after $<$ 100 ms exposure to the faces (Todorov, Pakrashi, & Oosterhof, 2009; Willis & Todorov, 2006), and can associate the affective value conveyed by an individual's behavior (moral or immoral) with their face in a few seconds (Bliss-Moreau, Barrett, & Wright, 2008; Todorov & Olson, 2008). Moreover, people can learn affective associations between as many as 500 unique faces and behaviors (Falvello et al., 2015), and can correctly retrieve such associations after days (e.g., Bliss-Moreau et al., 2008; Mattarozzi, Colonnello, Russo, & Todorov, 2018).

Although the robustness of the evaluation of others' characters is well established, it is unclear whether this striking ability is specific to learning about people or reflects a more general capacity to learn the positive or negative values of both human and non-human targets. Consistent with the latter possibility, Falvello et al. (2015) found that people could learn both the affective values of hundreds images depicting faces and places, and that this learning was comparable. In their paradigm, participants were briefly presented with images of faces paired with behavioral statements (moral, immoral, or neutral) and images of places paired with statements of valenced features of the place (positive, negative, or neutral). After seeing all the image-statement pairs, participants were asked to evaluate the goodness of a sample of the previously presented images. Findings from three experiments indicated that participants could easily create associations between the affective value conveyed by the statements and the images. Surprisingly, no differences were observed in learning the affective values of people and places. In fact, participants' evaluations were similarly influenced by the valence of the statements for both faces and places. The findings of Falvello et al. (2015) pose new questions about the underlying mechanisms of affective learning. Does learning about people tap on similar processes as learning about places? Moreover, ultimately, is there anything special about learning people's character?

The literature on the mechanisms underlying learning about other people from behavioral statements indicates multiple processes of associating the affective value conveyed by the statement and the face of the person performing the behavior. For instance, people make *trait inferences* from a behavioral statement and then associate the inferred

* Corresponding author at: Department of Brain and Behavioral Sciences, University of Pavia, Piazza Botta 6, 27100 Pavia, Italy.
*E-mail address:* chiara.ferrari@unipv.it (C. Ferrari).

https://doi.org/10.1016/j.actpsy.2020.103011
Received 13 March 2019; Received in revised form 20 December 2019; Accepted 10 January 2020
Available online 22 January 2020

trait with the face (Todorov & Uleman, 2002, 2003, 2004). For example, when people learn that someone found a large diamond ring and tried to locate the owner, they infer that the person performing the behavior is *honest* and this inference becomes associated with the person's face. *Trait inferences* rely on attributional processes (i.e., causal links resulting from cognitive processes), and they are sensitive to the characteristics of stimulus statements, so that they occur only when the statements reflect on the target (Todorov & Uleman, 2004). An alternative but not mutually exclusive process of associating the affective value conveyed by behavioral statements and faces is through affect-based mechanisms, as in the case of *evaluative inferences* (i.e., positive or negative affective reactions to behavioral statements that become associated with a person, e.g., Schneid, Carlston, & Skowronski, 2015). In this case, people can simply attribute a positive valence to the person who tried to locate the owner of the lost diamond ring, on the basis of the affective value conveyed by the statement but independently from the specific personality trait implied (e.g., honest). These affective mechanisms work irrespective of whether the perceiver considers the statements as accurate or inaccurate. In fact, the affective value elicited by the behavioral statement can be associated to the same extent with the face of the agent of the behavior or with the communicator of the behavior (Schneid et al., 2015). Similarly, processes related to classical conditioning (*CC*; e.g., Pavlov, 1927; for reviews, see Dickinson & Mackintosh, 1978; Pearce & Bouton, 2001) and evaluative conditioning (*EC*; e.g., Levey & Martin, 1975; for reviews, see De Houwer, Thomas, & Baeyens, 2001; Hofmann, De Houwer, Perugini, Baeyens, & Crombez, 2010) – that mainly rely on the spatiotemporal contiguity of a neutral and an affective stimulus – can also contribute to affective learning. Importantly, affect-based mechanisms (e.g., *CC* and *EC*) are at play when learning the affective value of objects: there is strong evidence that individuals are able to associate affective information with objects and that these affective associations influence subsequent perceptions (e.g., Ghuman & Bar, 2006). In line with this evidence, the extensive work on attitude formation (e.g., Olson & Fazio, 2001) shows that classical conditioning is a pervasive source of affective learning.

In the present study, we further explore the nature of the processes of learning the affective values of people and environments. Specifically, in four different experiments, we investigated whether general affect-based processes (in the form of the concurrent presentation of a neutral stimulus, such as a face, with a stimulus of known affective value, such as a valenced statement) are sufficient to account for learning the affective value of people and places.

## 2. Overview of the experiments

In Experiment 1, we adopted a procedure similar to the one of Falvello et al. (2015) in which participants were asked to evaluate the goodness of a sample of images, which were previously presented with valenced statements. In addition to the valence of the statements, we manipulated the relevance of the information conveyed by the statements to the images in order to create *relevant* (face image + behavioral statement or place image + place statement) and *irrelevant* trials (face image + place statement or place image + behavioral statement). We hypothesized that when image targets are places, participants would rate the images in accordance with the valence of the initial statement irrespective of their relevance (through general affect-based mechanisms). In turn, when targets are faces, we expected that participants' goodness ratings would be effectively modulated by valence only for relevant associations (possibly as a reflection of attributional processes that may be at least partially dissociable from general affect-based processes). In Experiment 1, we found that affective learning occurred exclusively on relevant trials. Surprisingly, this was not only the case for faces (in line with our predictions) but also for places. The latter finding appeared to contradict previous studies pointing to general affective-based processes as the mechanisms underpinning affective learning of inanimate objects. Therefore, to clarify this issue we conducted three additional experiments (Experiments 2, 3, and 4) in which, in contrast to Experiment 1 where the images were paired with a single statement (one-trial learning), we presented the images with multiple statements to test whether this repeated presentation can induce affect-based learning. This change was introduced, because previous studies indicate that general affect-based mechanisms require a rich experience in order to elicit an affective response toward an object and that associative learning requires a moderate number of experiences (Bliss-Moreau et al., 2008; Gawronski & Bodenhausen, 2006; for reviews, see Dickinson & Mackintosh, 1978; Pearce & Bouton, 2001). Across all three experiments, we found that affective associations with places were learned on both relevant and irrelevant trials, suggesting that learning about places taps on affect-based mechanisms. In contrast, affective associations with faces were learned only on relevant trials in Experiment 2 and 3 (in line with Experiment 1). In Experiment 4, in which we replaced the images of places with those of houses to ensure the generalizability of the observed effects, we found that affective learning of faces could also occur on irrelevant trials. However, the learning effect was larger for relevant than irrelevant trials. Our findings suggest that although both affective learning about people and non-human targets might rely on similar affect-based mechanisms, affective learning about people recruits person-attribution processes as well.

## 3. Experiment 1

In Experiment 1, we investigated whether affect-based mechanisms (the concurrent presentation of a face and a valenced statement) are sufficient to trigger affective learning. Moreover, we tested whether affect-based mechanisms contribute to affective learning of people and non-human objects (places) to a different extent. To do so, adopting a paradigm similar to that of Falvello et al. (2015), we created unique pairs of images (faces or places) and valenced (negative, neutral, or positive) statements. Half of the statements described social behaviors and half described features of places. We combined images and statements so that for some of the pairs the information conveyed by the statements referred to the target depicted in the image (relevant trials); while for the others it did not refer to the target (irrelevant trials). Specifically, in relevant pairs, photos of people were matched with behavioral statements and photos of places were matched with place statements. In irrelevant pairs, photos of people were presented with place statements and photos of places with behavioral statements, thus creating valenced image-statement pairs in which the information provided by the statement could not be properly related to the image. After being presented with all the image-statement pairs (each image-statement pair was presented once), participants were presented with images of faces and places alone (i.e., without the corresponding description), and asked to rate the goodness of each target. We expected that participants would rate the images in accordance with the valence of the initial statements, with those images associated with negative statements being evaluated more negatively than those associated with neutral and positive statements, and images associated with positive statements being evaluated more positively than those associated with neutral ones.

Moreover, we reasoned that if participants could learn affective associations on relevant and irrelevant trials, this would support the hypothesis that affective learning relies on simple affect-based mechanisms. In contrast, if participants learned affective associations only in relevant trials, this would support the hypothesis that affect-based mechanisms are not sufficient to account for affective learning, and that dissociable (possibly more complex) processes might take place. Importantly, we expected different learning processes as a function of the target stimulus (faces vs. places), with affective learning about people occurring only in relevant trials (in light of their cognitive/attributional nature) and affective learning about places occurring independently from the relevance of the statement (via general affect-based mechanisms).

### 3.1. Methods

#### 3.1.1. Participants

Twenty members of the Princeton community (7 males, mean age = 24.7 years) participated in exchange for 1 h of course credit or $12.

#### 3.1.2. Materials

Stimuli consisted of 360 pairs of images and statements. Half of the images pictured faces and half pictured outdoor scenes (all the images were presented in grayscale). Images of faces depicted un-posed faces of male and female individuals (90 males and 90 females) in a frontal view and were selected from the FERET Grayscale Database CD 1 (Phillips, Moon, Rizvi, & Rauss, 1999). Images of places depicted different outdoor scenes (e.g., mountains, streets, beaches, fields) and were selected from the collection of place images originally used by Konkle, Brady, Alvarez, and Oliva (2010). Statements were selected from a larger set used in a prior study (Falvello et al., 2015) and were either positively, negatively, or neutrally valenced, and described a person (for example, “He offered to help an elderly neighbor paint his house.”, “She didn't leave a tip at the restaurant.”, “He saw a funny movie several times.”) or a place (for example, “This ‘green’ home uses 85% solar power.”, “This beach often has blinding sandstorms.”, “This mountain is in Asia.”, see Falvello et al., 2015, for details on valence categorization of the statements). Fig. 1 shows examples of stimuli in the Experiment. Images and statements were paired in order to obtain relevant associations (images of people matched with behavioral statements and images of places paired with place statements) and irrelevant associations (images of people with place statements and images of places with behavioral statements). Images of people and places were randomly assigned to statements of different valence (positive, neutral, negative). Thirty unique image-statement associations were created for each category – stimulus displayed in the image (people and places), relevance (relevant and irrelevant), and valence (negative, neutral, and positive). When pictures of faces were presented, images were also balanced as a function of the gender of the individual pictured (fifteen photos of males and fifteen photos of females in each category).

#### 3.1.3. Procedure

The experiment was organized in two consecutive phases: an impression- formation phase and a rating task. In the impression-formation phase, participants were told that they would be viewing images of faces and places that would be paired with negative, neutral, or positive statements. They were instructed to read the statements and observe the images. Each trial consisted of the presentation of an image-statement pair for four and a half seconds, followed by a 500 ms fixation cross. The image-statement pairs were divided into three consecutively presented sets: an “initial set” composed of 60 images, an “intermediate set” composed of 240 images and a “final set” composed of 60 images. Within each set, the order of the image-statement pairs was randomized for each participant, but the order of the sets was fixed. Sets were presented one immediately after the other so that participants were not aware of the manipulation that was introduced in order to avoid recency and primacy effects in the rating task (see below).

After viewing all the images and statements, the rating task started. Participants were instructed to rate 240 photos (previously seen in the impression formation phase) on goodness on a 1–9 scale with 1 being “very bad”, 5 “neutral”, and 9 “very good”. It was also emphasized that they should rely on their intuitive, “gut” judgment, since there were no right or wrong answers. The same 240 place and face images were rated by all participants, and they were weighted equally for valence, relevance, stimulus, and gender (when applicable). Further, to avoid primacy and recency effects, images used in the rating belonged exclusively to the “intermediate set” of the impression-formation phase.

### 3.2. Results

We conducted a repeated-measures ANOVA (Bonferroni correction applied to all post-hoc comparisons) with valence of the statement (positive, negative, and neutral), relevance of the statement to the

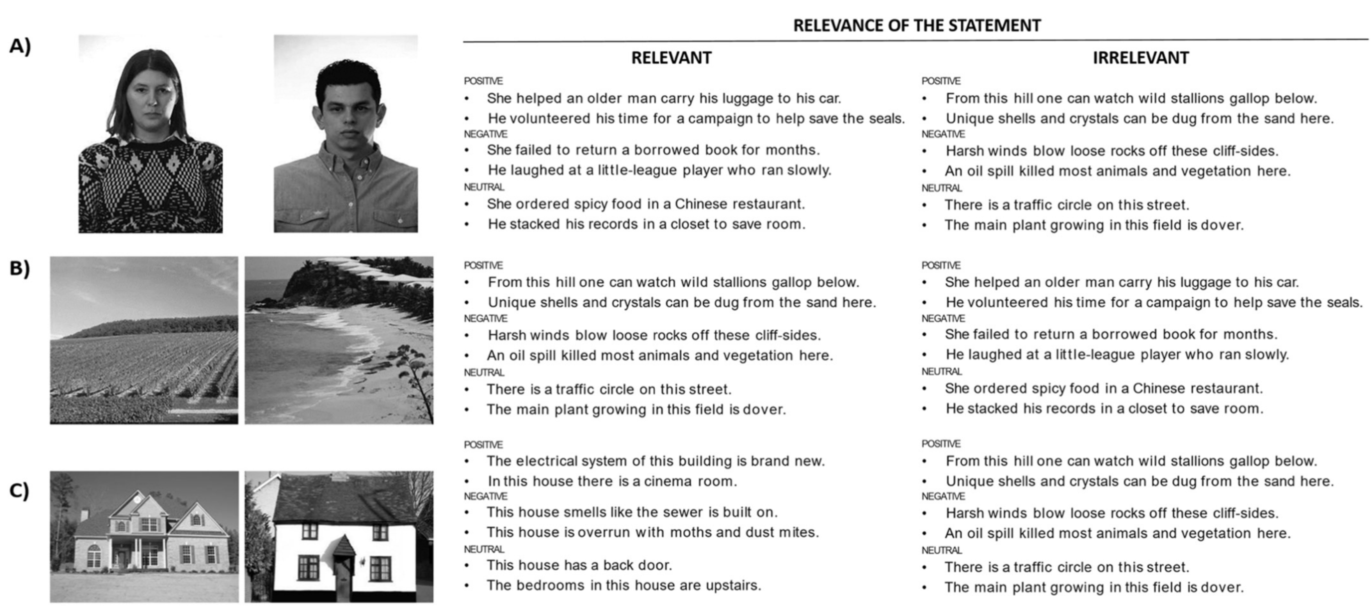


**Fig. 1.** Examples of images and statements used in the experiments as a function of the relevance of the information conveyed by the statement to the target image. A) Images depicting a male and a female individual and positive, negative and neutral statements referring to people's behaviors (face relevant statements) and positive, negative and neutral statements referring to places' valenced features (face irrelevant statements). B) Images depicting places (a field and a coast) and positive, negative and neutral statements referring to places' valenced features (place relevant statements) and positive, negative and neutral statements referring to people's behaviors (place irrelevant statements). C) Images depicting houses and positive, negative and neutral statements referring to houses' valenced features (house relevant statements) and positive, negative and neutral statements referring to places' valenced features (house irrelevant statements) (images of houses were used only in Experiment 4).

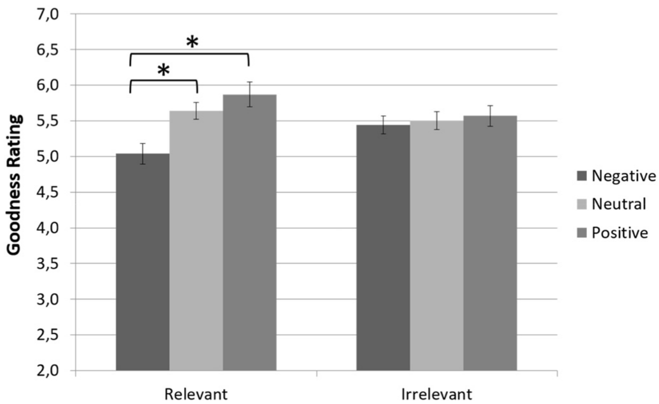


**Fig. 2.** Goodness judgments as a function of the valence (negative, neutral, and positive) and relevance of the statements (relevant and irrelevant) in Experiment 1. Affective learning occurred only when relevant statements were provided (irrespective of the stimulus evaluated). Standard errors are standard errors of the mean. The asterisks represent significant differences between experimental conditions.

image (relevant and irrelevant), and stimulus pictured in the image (people and places) on goodness ratings.

Neither the main effect of relevance nor the main effect of stimulus reached significance, $F(1,19) < 1$, $p = .79$, and $F(1,19) = 2.19$, $p = .16$, respectively. A significant main effect of valence, $F(2,38) = 18.66$, $p < .001$, $\eta_p 2 = 0.50$, was found, qualified by a significant interaction with relevance, $F(2,38) = 13.17$, $p < .001$, $\eta_p 2 = 0.41$ (see Fig. 2). To clarify the interaction, we conducted an analysis of the main effect of valence separately for relevant and irrelevant trials.

In relevant trials, the main effect of valence was significant, $F(2,38) = 21.01$, $p < .001$, $\eta_p 2 = 0.53$. Post-hoc comparisons showed that faces and places paired with negative statements were rated as worse than those paired with neutral statements ($p < .001$) and those paired with positive statements ($p < .001$). No significant difference was found between ratings of images paired with positive and neutral statements ($p = .39$), although the difference was in the predicted direction. In the irrelevant condition, the valence of the statements did not affect goodness ratings, $F(2,38) = 1.45$, $p = .25$, suggesting that when the information provided by the statement was not relevant to the target of the evaluation the affective learning did not occur. No other two-way and the three-way interaction effects reached significance ($ps > .18$).

### 3.3. Discussion

Participants could effectively learn the affective values of both faces and places as long as statements conveyed information relevant to the images. In fact, when targets were matched with irrelevant statements (i.e., an image of a face paired with a place statement or an image of a place paired with a behavioral statement), judgments were not affected by the valence of the statement. Critically, this was the case for both faces and places. The fact that affective learning of people occurs in a relevance-specific fashion fits well with previous studies pointing to the importance of cognitive/attributional mechanisms underpinning this process (e.g., Bliss-Moreau et al., 2008; Todorov & Uleman, 2004). However, the same pattern of results for places is inconsistent with our predictions based on extensive evidence indicating that affect-based mechanisms account for learning of the affective value of inanimate objects (e.g., Ghuman & Bar, 2006; Pearce & Bouton, 2001). Moreover, this result indicates that learning of the affective value of people and places depends on similar processes.

However, it is important to note that while in our experiment participants had a single chance to experience the association between the neutral and the affective stimulus (each image-statement pair was indeed presented once), previous studies indicate that more experiences are needed to trigger general affect-based learning (e.g., Dickinson & Mackintosh, 1978; Ghuman & Bar, 2006; Pearce & Bouton, 2001). Therefore, it is possible that a single experience with the source of the affective value was not sufficient to elicit affect-based mechanisms, preventing participants to learn from irrelevant associations. In Experiments 2 to 4, we investigated whether affect-based mechanisms might account for affective learning of faces and places in the case of multiple experiences with the affective stimulus.

## 4. Experiment 2

In Experiment 2, we asked a new group of participants to evaluate the goodness of images of faces and places previously presented with relevant and irrelevant valenced statements. In contrast to Experiment 1, here images were paired with two statements of the same valence. Consistent with our original predictions, we expected affective learning of people to occur only in relevant trials and affective learning of places to occur irrespective of the relevance of the statement.

### 4.1. Methods

#### 4.1.1. Participants

Thirty members of the Princeton community (10 males, mean age = 19.6 years) participated in exchange for 1 h of course credit or $12.

#### 4.1.2. Materials

Stimuli consisted of 240 image-statement pairs selected from those used in Experiment 1. The number of image-statement pairs was reduced to 240 (in Experiment 1 we used 360 pairs), because here each image-statement pair was repeated twice for a total number of 480 trials. Image-statement pairs were equally weighted for the valence of the statement (positive, negative, and neutral), the relevance of the statement to the target (relevant and irrelevant), the type of stimulus displayed in the image (people and places), and gender (when applicable). Furthermore, half of the images were always presented with the same statements and half were presented with two different statements. When an image was paired with different statements, the statements always conveyed the same valence.

#### 4.1.3. Procedure

The impression formation task was similar to that in Experiment 1 with the exception that each image was repeated twice. After viewing all the images and statements, participants were asked to rate all of the 240 pictures on goodness on a scale from 1 to 9. The images were weighted equally for valence (positive, negative, and neutral), relevance (relevant and irrelevant), stimulus (people and places), gender (when applicable), and type of the statement (same statement and different statements). Concerning the latter manipulation, although we distinguished between targets matched with the same repeated statement and with different statements, we failed to find a clear pattern across Experiments 2 and 3 and importantly the type of the statement did not qualify our main findings. Therefore, we collapsed data across this factor in the first analysis reported below and, in a following analysis, we analyzed goodness ratings separately for targets matched with the same repeated statements and with different statements.

### 4.2. Results

We run a repeated-measures ANOVA (Bonferroni correction applied to all post-hoc comparisons) with valence of the statements (positive, negative, and neutral), relevance of the statement to the image (relevant and irrelevant), and stimulus pictured in the image (people and places) on goodness ratings.

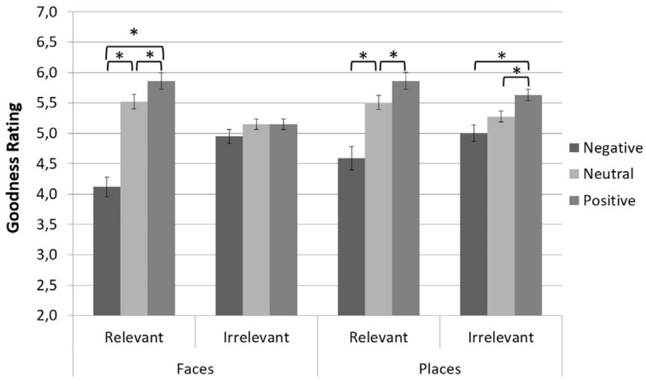


**Fig. 3.** Goodness judgments as a function of the valence of the statements (negative, neutral, and positive), the relevance of the statements (relevant and irrelevant) and the type of stimulus (faces and places) in Experiment 2. For faces, affective learning occurred only when relevant statements were provided. In contrast, for places, affective learning occurred both when relevant and irrelevant statements were provided. Standard errors are standard errors of the mean. Asterisks represent significant differences between experimental conditions.

Participants learned about both human and non-human objects, as indicated by the main effect of valence, $F(2,58) = 33.36$, $p < .001$, $\eta_p2 = 0.54$ (no other main effects were significant: relevance, $F(1,29) = 2.73$, $p = .11$, and stimulus, $F(1,29) = 2.27$, $p = .14$). Similarly to Experiment 1, the learning process was affected by the relevance of the statement associated with the images, as indicated by the significant interaction of valence by relevance, $F(2,58) = 29.73$, $p < .001$, $\eta_p2 = 0.51$. However, this interaction effect was further qualified by the three-way interaction of valence by relevance by stimulus, $F(2,58) = 9.87$, $p = .001$, $\eta_p2 = 0.25$. We investigated this three-way interaction by analyzing the effects for faces and places separately (see Fig. 3).

For judgments of faces, both the main effect of valence, $F(2,58) = 33.81$, $p < .001$, $\eta_p2 = 0.54$ (but not of relevance, $F(1,29) = 2.98$, $p = .10$), and the interaction of valence by relevance reached significance, $F(2,58) = 34.39$, $p < .001$, $\eta_p2 = 0.54$. Replicating Experiment 1, the interaction indicated that participants formed affective associations with faces when the information was relevant to the face, $F(2,58) = 44.39$, $p < .001$, $\eta_p2 = 0.61$ (all post-doc comparisons being significant, all *ps* $< .028$), but not when it was irrelevant, $F(2,58) = 2.12$, $p = .13$.

For judgments of places, both the main effect of valence, $F(2,58) = 24.48$, $p < .001$, $\eta_p2 = 0.46$, (but not the main effect of relevance, $F(1,29) < 1$, $p = .64$) and the interaction of valence by relevance, $F(2,58) = 9.44$, $p = .001$, $\eta_p2 = 0.25$, reached significance. The interaction indicated that although participants could learn the affective values of places both in the relevant, $F(2,58) = 22.36$, $p < .001$, $\eta_p2 = 0.44$, and in the irrelevant condition, $F(2,58) = 14.53$, $p < .001$, $\eta_p2 = 0.33$, the effect was slightly stronger for the relevant condition (see Table 1).

Consistent with the above analyses, analyses on goodness ratings of places and faces given in irrelevant trials (repeated-measures ANOVA with stimulus and relevance as within-subjects factors) showed a significant interaction effect of valence by stimulus, $F(2,58) = 9.79$, $p < .001$, $\eta_p2 = 0.25$. This interaction indicated that affective learning occurred in the case of places, $F(2,58) = 14.53$, $p < .001$, $\eta_p2 = 0.33$ (with images of places associated with negative statements rated as more negative than those associated with neutral ones, $p = .06$, and positive statements, $p < .001$, and images associated with positive statements rated as more positive than those associated with neutral statements, $p = .002$), but not in the case of faces, $F(2,58) = 2.12$, $p = .13$.

Lastly, to explore whether varying the type of behavioral information from which perceivers extracted the affective value would influence the way in which individuals learn about others, we replicated our first analysis (ANOVA with as independent factors valence of the statement, relevance of the statement, and stimulus) on the goodness ratings, separately for each type of statement repetition (repetition of same statement and presentation of different statements of the same valence). For the sake of brevity, we report here only relevant statistics, while detailed analysis relative to this manipulation is described in the Supplementary Material.

**Same statement repetition**: The analysis revealed a significant interaction effect valence by relevance by stimulus, $F(2,58) = 14.47$, $p < .001$, $\eta_p2 = 0.33$, indicating that affective learning in the relevant and irrelevant condition occurred differently for faces and places. Specifically, affective learning of faces occurred only for relevant statements, $F(2,58) = 42.65$, $p < .001$, $\eta_p2 = 0.60$ (affective learning for irrelevant statements, $F(2,58) = 1.06$, $p = .35$). In turn, for places, the affective learning occurred for both relevant, $F(2,58) = 26.25$, $p < .001$, $\eta_p2 = 0.48$, and (to a slightly lower extent) for irrelevant trials, $F(2,58) = 23.15$, $p < .001$, $\eta_p2 = 0.44$.

**Different statement repetition**: The analysis revealed a significant effect relevance by valence, $F(2,58) = 21.68$, $p < .001$, $\eta_p2 = 0.43$, indicating that affective learning occurred in relevant trials, $F(2,58) = 27.07$, $p < .001$, $\eta_p2 = 0.48$, but not in irrelevant trials, $F(2,58) = 1.45$, $p = .24$ (this effect was consistent for both faces and places).

### 4.3. Discussion

Experiment 2 showed that learning the affective value of places occurred when images were presented with two statements describing both relevant and irrelevant information. In contrast, learning the affective value of people took place only when the statements described social behaviors, but not when the statements were irrelevant. These findings suggest that, over multiple experiences, the processes underlying the learning of the affective value of people and places could be

**Table 1**
Effect size (Cohen's *d*), mean difference, and 95% confidence intervals of affective learning across Experiments 2, 3, and 4. Cohen's *d* and mean difference (95% CI) of goodness ratings between images associated with positive and negative statements as a function of the stimulus type (faces/non-human targets) and the statement relevance (relevant/irrelevant) in Experiment 2, 3, and 4.

| | | Human targets | | Non-human targets | |
|---|---|---|---|---|---|
| | | Relevant | Irrelevant | Relevant | Irrelevant |
| Experiment 2 | Cohen's d | 1.30 | 0.27 | 0.92 | 0.81 |
| | Mean difference (95% CI) | 1.74 (1.21–2.24) | 0.20 (−0.48–0.78 >) | 1.28 (0.76–1.79) | 0.63 (0.34–0.92) |
| Experiment 3 | Cohen's d | 1.27 | 0.52 | 1.26 | 1.09 |
| | Mean difference (95% CI) | 1.82 (1.06–2.58) | 0.22 (−0.01–0.44) | 1.49 (0.86–2.12) | 1.39 (0.71 > –2.07) |
| Experiment 4 | Cohen's d | 1.01 | 0.60 | 1.02 | 0.75 |
| | Mean difference (95% CI) | 1.46 (1.09–1.83) | 0.32 (0.18–0.46) | 0.76 (0.57–0.95) | 0.47 (0.31–0.64) |

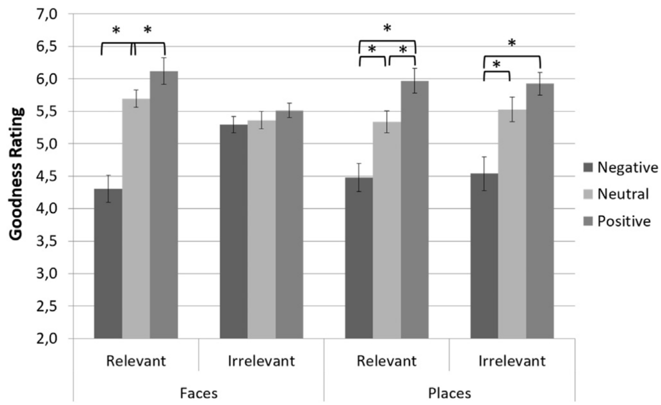


**Fig. 4.** Goodness judgments as a function of the valence of the statements (negative, neutral, and positive), the relevance of the statements (relevant and irrelevant), and the stimulus (faces and places) in Experiment 3. For faces, affective learning occurred only when relevant statements were provided. In contrast, for places affective learning occurred both when relevant and irrelevant statements were available. Standard errors are standard errors of the mean. Asterisks represent significant differences between experimental conditions.

different. The simple temporal and spatial contiguity between a neutral and an affective stimulus is sufficient to account for affective learning about objects but not people.

## 5. Experiment 3

The objective of Experiment 3 was to replicate the findings of Experiment 2, using more experiences with the affective stimulus (image-statements pairs were repeated three times in Experiment 3). Moreover, one critical issue concerning the previous experiments is that image-statement pairs for both faces and places were intermixed in the same block, possibly emphasizing differences (rather than similarities) between learning of the two categories. Therefore, in Experiment 3, we adopted a between-subjects design in which participants were presented either with images of faces or images of places, to verify that differences in the affective learning of people and places found in Experiment 2 generalize to separate evaluations of human and non-human targets.

### 5.1. Methods

#### 5.1.1. Participants

Thirty-two members of the Princeton community (19 males, mean age = 19.5 years) participated in exchange for 1 h of course credit or $12.

#### 5.1.2. Materials and procedure

Pictures of people and places, as well as statements, were selected from those of the previous experiments. In order to rule out that the effects observed in Experiment 2 might be specific to the within-subjects design, we decided to adopt a between-subjects design where two different groups of participants were exposed to either images of faces or places. In the impression formation phase, the number of image-statement pairs was further reduced to 180 (in Experiment 2 we used 240 pairs) because here each image-statement pair was repeated three times for a total number of 540 trials. Image-statement pairs (either referring to places or faces) were equally weighted for the valence of the statement (positive, negative, and neutral), the relevance of the statement to the target (relevant and irrelevant), and gender (when applicable). Furthermore, half of the images were always presented with the same statements and half were presented with three different statements. When an image was paired with different statements, the statements always conveyed the same valence. Similarly to Experiment 2, this manipulation did not qualify our main findings; therefore, we collapsed data across this factor in the first analysis reported below. However, we also report a following analysis in which we analyzed goodness ratings separately for the two types of statements. Image-statement pairs were organized in three consecutively presented sets (an "initial set" composed of 30 image-statement pairs, an "intermediate set" composed of 120 image-statement pairs and a "final set" composed of 30 image-statement pairs). After viewing all the images and statements, participants were asked to rate the same 120 pictures of either faces or places (depending on the task to which they were assigned) on goodness on a scale from 1 to 9. To avoid primacy and recency effects, images used in the rating task belonged exclusively to the "intermediate set" of the impression formation phase.

### 5.2. Results

We run a mixed-measures ANOVA (Bonferroni correction applied to all post-hoc comparisons) with valence of the statements (positive, negative, and neutral), relevance of the statement to the image (relevant and irrelevant) as within-subjects factors, and type of stimulus pictured in the image (people and places) as a between-subjects factor on goodness ratings.

The analyses showed that participants learned about both human and non-human objects, as indicated by the main effect of valence, $F(2,60) = 38.77$, $p < .001$, $\eta_p2 = 0.56$ (no other main effects were significant: relevance, $F(1,30) < 1$, $p = .51$; stimulus, $F(1,30) < 1$, $p = .60$). In line with Experiment 2, the interactions of valence by relevance, $F(2,60) = 15.31$, $p < .001$, $\eta_p2 = 0.34$, and valence by relevance by stimulus, $F(2,60) = 14.56$, $p < .001$, $\eta_p2 = 0.33$, were significant. Other two-way and three-way interactions did not reach significance ($ps > 0.34$). We investigated the three-way interaction by analyzing the interaction effect of valence by relevance for faces and places separately (see Fig. 4).

For faces, replicating the findings of Experiment 1 and Experiment 2, the significant main effect of valence, $F(2,30) = 22.78$, $p < .001$, $\eta_p2 = 0.60$ (the main effect of relevance was not significant, $p = .77$) was qualified by the interaction of valence by relevance, $F(2,30) = 22.59$, $p < .001$, $\eta_p2 = 0.60$, showing that participants learned about the affective value of people in the relevant, $F(2,30) = 25.23$, $p < .001$, $\eta_p2 = 0.63$, but not in the irrelevant condition, $F(2,30) = 2.76$, $p = .08$.

In contrast, for places, only the main effect of valence, $F(2,30) = 18.73$, $p < .001$, $\eta_p2 = 0.56$, was significant, indicating that participants learned the affective values of places to the same extent whether the places were presented with relevant or irrelevant statements. Neither the main effect of relevance nor the interaction of relevance by valence reached significance ($ps > 0.46$).

Similarly to Experiment 2, the analyses on goodness ratings of places and faces given in irrelevant trials showed a significant interaction effect of valence by stimulus, $F(2,60) = 9.79$, $p < .001$, $\eta_p2 = 0.23$, indicating that affective learning occurred when targets were places, $F(2,30) = 13.56$, $p < .001$, $\eta_p2 = 0.48$. Specifically, images of places associated with negative statements were rated as more negative than those associated with neutral, $p = .016$, and positive statements, $p = .002$. No significant difference was found between ratings of images paired with positive and neutral statements ($p = .12$), although the difference was in the predicted direction. Critically, the main effect of valence did not reach significance in the case of faces, $F(2,30) = 2.76$, $p = .08$.

Lastly, we replicated our first analysis (ANOVA with as independent factors valence of the statement, relevance of the statement, and stimulus) on the goodness ratings, separately for each type of statement repetition (repetition of the same statement and presentation of different statements of the same valence). We report here only relevant statistics, while detailed analysis relative to this manipulation is described in the Supplementary Material.

**Same statement repetition**: The analysis revealed a significant effect of valence by relevance by stimulus, $F(2,60) = 13.33$, $p < .001$, $\eta_p2 = 0.31$, indicating that affective learning in relevant and irrelevant trials occurred differently for faces and places. Although the affective associations were learned to a bigger extent for relevant ($F(2,30) = 22.48$, $p < .001$, $\eta_p2 = 0.60$, for faces and $F(2,30) = 15.23$, $p < .001$, $\eta_p2 = 0.50$, for places) than irrelevant trials ($F(2,30) = 5.48$, $p = .009$, $\eta_p2 = 0.27$, for faces and $F(2,30) = 8.53$, $p = .001$, $\eta_p2 = 0.36$, for places), the interaction of relevance by valence was much smaller in magnitude for places ($F(2,30) = 3.53$, $p = .042$, $\eta_p2 = 0.19$) than for faces ($F(2,30) = 24.09$, $p < .001$, $\eta_p2 = 0.62$).

**Different statement repetition**: The analysis revealed a significant effect of valence by relevance by stimulus, $F(2,60) = 9.02$, $p < .001$, $\eta_p2 = 0.23$, indicating that for faces, affective learning occurred in relevant trials, $F(2,30) = 21.27$, $p < .001$, $\eta_p2 = 0.59$, but not in irrelevant trials, $F(2,30) = 2.51$, $p = .09$. For places, only the main effect of valence, $F(2,30) = 18.57$, $p < .001$, $\eta_p2 = 0.55$, reached significance, indicating that affective learning occurred similarly in relevant and irrelevant trials.

### 5.3. Discussion

Experiment 3 confirmed the findings of Experiment 2, showing that the affective value of places could be learned in both relevant and irrelevant trials, while the affective value of people only in relevant trials. These findings provide additional evidence about possible differences in learning the affective value of people and places, even in the case of separate evaluations of human and non-human targets.

## 6. Experiment 4

Experiment 4 aimed to overcome potential limitations of the previous experiments and increase the generalizability of the observed effects. First, in the previous experiments, we compared images of places to those of faces because, similarly to faces, places have a remarkable relevance to people's lives and are often targets of affective learning. However, faces and places differ in their visual characteristics, which might affect the way people recognize and learn about them. For instance, while faces are represented holistically (e.g., Maurer, Le Grand, & Mondloch, 2002), the perception of places rely on more featural properties. In the fourth experiment, we wanted to rule out the possibility that our results were driven by visual features and that our findings were specific to place images. Therefore, we tested whether we can replicate our prior findings using images of houses. Moreover, in the first three experiments, to create perfectly counterbalanced irrelevant associations, we paired images of faces with place-related statements and images of places with behavioral statements. However, it is possible that in place-irrelevant trials, participants imagined themselves in the place shown on the image engaging in the behavior described in the statement. Such a strategy might have favored affective learning of places (but not faces) in irrelevant trials, as we observed in the previous experiments. Therefore, in Experiment 4, we a) replaced images of places with images of houses and b) paired both face and house images with irrelevant place-statements to create incongruent trials. Images of houses were chosen as new control stimuli because, similarly to faces, they are represented holistically (Bruce, Doyle, Dench, & Burton, 1991) and have been used as control stimuli for faces in many previous studies (e.g., Pitcher, Walsh, Yovel, & Duchaine, 2007; Shigihara & Zeki, 2014; Summerfield, Egner, Mangels, & Hirsch, 2005). In addition, we asked two independent groups of individuals to rate the images used in the experiment on valence ($N = 32$), and the statements on arousal ($N = 32$). These additional rating tasks were introduced to assure that all images were intrinsically neutral in valence, that the emotional value of the images was conveyed exclusively by the statements, and that all the statements were equally arousing. Lastly, in Experiment 4, we enlarged our sample size ($N = 60$) to increase statistical power. A power analysis using the *R pwr* library (Champely et al., 2015) indicated that 45 participants would afford a power of 0.8 to detect the valence by relevance by stimulus interaction. The power analysis was based on the effect size of the 3-way interaction found in Experiment 3 ($\eta_p2 = 0.33$).

### 6.1. Methods

#### 6.1.1. Participants

Sixty members of the Princeton community (21 males, mean age = 19.4 years) participated in exchange for 1 h of course credit or $12.

#### 6.1.2. Materials

Stimuli consisted of pictures of people and houses (see Fig. 1 for examples of the stimuli used). Photos of people were selected from those of the previous experiments and photos of houses were selected from the web using Google Image Search. We edited the house images such that they were all grayscale, and have the same size and resolution as those of people images. We used valenced behavioral, place and house statements. Behavioral and place statements (referring to natural scenery such as mountains, beaches, coasts, fields, and gardens but not to houses or buildings) were the same as the ones used in the previous experiments. House statements were generated in our lab and specifically referred to valenced features of houses or residential buildings (for example, "This building houses a library with a very rich collection." for the positive condition; "This house faces Southwest." for the neutral condition; "This house has many roof leaks." for the negative condition). The stimuli were combined to obtain the following four categories: 1) face image + behavioral statement (person relevant associations), 2) house image + house statement (house relevant associations), 3) face image + place statement (people irrelevant associations), and 4) house image + place statement (house irrelevant associations).[1] In order to ensure that the images used in Experiment 4 were neutral in valence across all conditions (i.e., irrespective of the statements matched with them), all the images used in the experiment were rated on valence by an independent group of online participants ($N = 32$, 16 males, mean age = 34.5 years). Participants were sequentially presented with images of either faces ($N = 16$) or houses ($N = 16$) without corresponding statements and asked to evaluate the valence of each image on a 1–9 valence scale (1 = very negative, 9 = very positive). A 2 [stimulus type: face/house] × 3 [statement valence: positive/neutral/negative] ANOVA revealed that overall the images alone were neutral in valence (mean = 5.2, SD = 1.0) and that valence evaluations did not differ depending on the stimulus pictured in the images and the valence of the statement to which they would be matched in the main task (main effect of stimulus type, $p = .52$, main effect of statement valence, $p = .12$, interaction effect of stimulus type by statement valence, $p = .13$).

In order to make sure that positive and negative statements were similarly effective in triggering affecting learning (compared to neural statements), we ran an additional online rating task on arousal. The arousal rating task ($N = 32$; 17 males, mean age = 36.4 years) consisted of the consecutive presentations of the statements used in Experiment 4 without the corresponding images that each participant had to rate on arousal (1–9 scale). To avoid participant fatigue, each participant rated only 120 statements out of the 240 total statements. A 2 [statement target: people/house] × 3 [statement valence: positive/neutral/negative] ANOVA on the arousal ratings revealed that

[1] Due to technical reasons, one house image was not displayed correctly. Therefore, the corresponding trial was not included in the data analysis.

statements referring to people (mean = 4.90, SD = 1.21) received higher arousal evaluation compared to those referring to houses (mean = 4.22, SD = 1.20, $p < .001$) and that arousal evaluations depended on the valence they conveyed ($p < .001$). Specifically, arousal ratings were higher when statements conveyed a negative or positive valence compared to a neutral valence (all $ps < 0.001$). Importantly, no difference was observed in the arousal ratings between positive and negative statements ($p = .55$).

#### 6.1.3. Procedure

In the impression formation phase, the number of image-statement pairs was set to 240 (sixty unique image-statement associations were created for each category), all repeated two times for a total number of 480 trials. Image-statement pairs (either referring to houses or faces) were equally weighted for the valence of the statement (positive, negative, and neutral), the relevance of the statement to the target (relevant and irrelevant), and gender (when applicable). After viewing all the images and statements, participants were asked to rate all the pictures on goodness on a scale from 1 to 9.

### 6.2. Results

We run a repeated-measures ANOVA (Bonferroni correction applied to all post-hoc comparisons) with valence of the statements (positive, negative, and neutral), relevance of the statement to the image (relevant and irrelevant), and stimulus pictured in the image (people and houses) as within-subjects factors on goodness ratings.

The analyses showed that participants learned about both human and non-human objects, as indicated by the main effect of valence, $F(2,118) = 87.16$, $p < .001$, $\eta_p 2 = 0.60$. The main effect of relevance, $F(1,59) = 30.50$, $p < .001$, $\eta_p 2 = 0.34$, and stimulus, $F(1,59) = 8.35$, $p = .005$, $\eta_p 2 = 0.13$, were also significant. The interactions of stimulus by relevance, $F(1,59) = 5.91$, $p = .018$, $\eta p2 = 0.09$, valence by stimulus, $F(2,118) = 8.85$, $p < .001$, $\eta_p 2 = 0.13$, valence by relevance, $F(2,118) = 46.54$, $p < .001$, $\eta_p 2 = 0.44$, and, critically, valence by relevance by stimulus, $F(2,118) = 10.44$, $p < .001$, $\eta_p 2 = 0.15$, were significant. We investigated the three-way interaction by analyzing the interaction effect of valence by relevance for faces and houses separately (see Fig. 5).

For faces, the significant main effect of valence, $F(2,118) = 56.99$, $p < .001$, $\eta_p 2 = 0.49$ (the main effect of relevance, $F(1,59) = 3.49$, $p = .07$, $\eta_p 2 = 0.06$ was not significant) was qualified by the interaction of valence by relevance, $F(2,118) = 34.23$, $p < .001$, $\eta_p 2 = 0.37$. Although participants learned about the affective value of people both in relevant, $F(2,118) = 56.51$, $p < .001$, $\eta_p 2 = 0.49$, and irrelevant trials, $F(2,118) = 10.66$, $p < .001$, $\eta p2 = 0.15$, affective learning was stronger in relevant than in irrelevant trials (see Table 1).

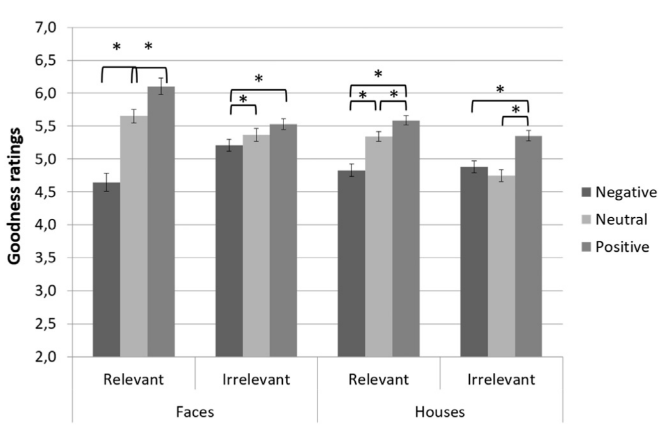


**Fig. 5.** Goodness judgments as a function of the valence of the statements (negative, neutral, and positive), the relevance of the statements (relevant and irrelevant), and the stimulus (faces and houses) in Experiment 4. Affective learning of people and places occurred both when relevant and irrelevant statements were provided. Standard errors are standard errors of the mean. Asterisks represent significant differences between experimental conditions.

For houses, both the main effect of valence, $F(2,118) = 51.59$, $p < .001$, $\eta p2 = 0.47$, and the main effect of relevance, $F(1,59) = 43.78$, $p < .001$, $\eta_p 2 = 0.43$, were significant. The interaction of valence by relevance reached significance, $F(2,118) = 15.12$, $p < .001$, $\eta_p 2 = 0.20$, indicating that participants were able to learn the affective value of houses both in relevant ($F(2,118) = 39.15$, $p < .001$, $\eta_p 2 = 0.40$) and irrelevant trials ($F(2,118) = 29.21$, $p < .001$, $\eta_p 2 = 0.33$), but with a slightly stronger effect in relevant than in irrelevant trials (see Table 1).

Similarly to the previous experiments, the analysis on goodness ratings of houses and faces given in irrelevant trials showed a significant interaction effect of valence by stimulus, $F(2,118) = 8.78$, $p < .001$, $\eta_p 2 = 0.130$, confirming that affective learning occurred differently for the two types of stimuli. Indeed, although affective learning occurred both when targets were faces, $F(2,118) = 10.66$ $p < .001$, $\eta_p 2 = 0.15$, and houses, $F(2,118) = 29.21$, $p < .001$, $\eta_p 2 = 0.33$, affective learning in irrelevant trials was slightly bigger for houses than faces (see Table 1). Moreover, ratings of faces associated with negative statements were significantly lower than those associated with positive ($p < .001$) and neutral statements ($p = .028$), but ratings of faces associated with neutral and positive statements did not differ, $p = .13$. In turn, for houses, the pattern was reversed, with ratings of images associated with positive statements rated as more positive than those associated with neutral ($p < .001$) and negative statements ($p < .001$), and no difference between houses associated with negative and neutral statements ($p = .33$) (this is possibly due to the particularly low goodness ratings given in the neutral condition, see Fig. 5).

## 7. General discussion

In four experiments, participants were presented with faces and places (or houses) paired with statements varying in valence and were subsequently asked to judge the goodness of each individual or place presented alone. Consistent with previous evidence (Bliss-Moreau et al., 2008; Falvello et al., 2015; Schneid et al., 2015; Todorov & Uleman, 2002, 2003), we found that participants were able to form and retrieve valence associations with faces, even after learning a single or few pieces of information. Furthermore, consistent with Falvello et al. (2015), participants were also able to learn the affective values of places and houses.

The present study demonstrates both similarities and differences in the affective learning about people and non-human targets as a function of the relevance of the statement to which they were paired and as a function of the number of experiences (single vs. multiple experiences) with the source of the affective value. Indeed, when target images and valenced statements were presented only once each (Experiment 1), affective learning occurred similarly for faces and places. Participants learned associations between images of faces or places and valenced statements only when statements described relevant affective information. When pictures of faces and places were associated with irrelevant information, the goodness ratings were not affected by the valence of the statements.

However, when statements were repeatedly presented (Experiments 2, 3, and 4), learning the affective values of people and places seemed to follow different routes. For faces, we observed the same pattern of results as in Experiment 1, with participants still learning associations between faces and valenced statements only when the faces were matched with statements that described social behaviors (relevant condition), but not when pictures of faces were associated with irrelevant information (place statements). In turn, for places, affective learning occurred irrespectively of the relevance of the statement to the image. Critically, the findings of Experiment 4 (in which we increased the sample size to perform more powerful analyses and replaced the images of places with those of houses to test the generalizability of our findings) showed that participants could learn affective associations of faces also in the irrelevant condition. This latter finding, inconsistent with our hypothesis, suggests that the temporal and spatial contiguity between an affective statement and an image might be sufficient to trigger an effective emotional association of faces with irrelevant valenced statements. We are not the first to report that valence-based processes characterize affective learning about people. Indeed, these processes are evident in spontaneous *trait transference*, a phenomenon demonstrated by Carlston, Skowronski, and colleagues (e.g., Carlston & Skowronski, 2005; McCarthy, Wells, Skowronski, & Carlston, 2018; Skowronski, Carlston, Mae, & Crawford, 1998). They showed that individuals attributed a trait conveyed by a behavior description to an agent, although the behavior was not enacted by the agent, possibly relying on associative mechanisms. In line with this finding, *evaluative inferences* occurring when participants associate a target face with the affective meaning (valence) but not the specific personality trait conveyed by a behavioral statement seem to rely on associative processes (e.g., Gawronski & Bodenhausen, 2006; Schneid et al., 2015). Critically, *transference effects* have been observed about agents paired with social behaviors, namely behaviors that are (at least minimally) relevant to the target of the evaluation. The present findings demonstrate that information that cannot be related to the target person at all might still trigger affective learning. Our findings also show that this general valence effect is far more consistent and stronger for non-human than human targets. Indeed, affective learning about faces in the irrelevant condition was only significant in Experiment 4, but not in Experiment 2 and 3. The latter difference between the experiments can be attributed to the lower statistical power of Experiments 2 and 3. However, across all three experiments (Experiments 2–4), the affective value of people had a much stronger effect in the relevant than the irrelevant condition (see Table 1). Importantly, as revealed by no cross-condition difference in arousal ratings of the statements, it is highly likely that the affective learning observed across the studies reflects the affective value of the statements rather than their arousal level.

The last three experiments (Experiments 2–4) provide convergent evidence that the extent to which participants learned affective associations in irrelevant trials was smaller for face images than for their non-human counterparts (see Table 1 and analyses on goodness ratings on irrelevant trials). These results possibly indicate that affect-based mechanisms might be less effective in triggering affective associations about people. This is consistent with the conclusions of the study of Bliss-Moreau et al. (2008), who asked participants to categorize faces previously associated with valenced verbal statements as positive, negative, or neutral. Statements to which faces were paired varied not only as a function of valence (positive, negative, and neutral) but also as a function of their "social content". That is, faces could be associated with valenced *social* statements (describing positive, negative, and neutral actions directed toward people, e.g., "Celebrated a child's birthday") and valenced *non-social* statements (describing positive, negative, and neutral actions directed to non-human targets, e.g., "Smelled fresh baked cookies"). Consistent with our findings, the affective categorization of faces reflected the valence of the statement with which the face was paired (e.g., faces associated with negative statements were categorized as negative), but this affective learning occurred to a bigger extent when faces were associated with *social* than with *non-social* statements (particularly for negative affective associations). In line with these prior findings, our results suggest that learning about others, although relying on shallow valence inferences to some extent, might not be completely explained by a spatiotemporal contiguity of an affective and a neutral stimulus alone. Accordingly, Todorov and Uleman (2003) showed that reading that an individual behaved in a specific moral way goes beyond the valence association between that behavior and the individual and elicits a specific personality trait inference that is embedded in the individual's representation.

With respect to shallow valence inferences, while the present research shows the role of general affect-based learning, it does not address whether the strength of the affective learning about people and non-human targets in the irrelevant condition reflects different subprocesses (associative vs. attributional) or a single process taking place to a different extent (e.g., a single associative/affect-based process). The idea of a single affect-based process is consistent with a recent study proposing that the same automatic associative links underlying associative phenomena in person evaluation (e.g., *transference effects*) can also explain inferential mechanisms of person attribution. In this view, differences between association and attribution depend on the level of activations of the association between the presented stimuli rather than on qualitatively different mechanisms (Orghian, Garcia-Marques, Uleman, & Heinke, 2015). Future studies need to test these different possibilities.[2]

Nonetheless, our studies consistently show that the simultaneous presentation of a non-human target (places/houses) and multiple valenced, but irrelevant statements is sufficient to trigger affective associations. This finding suggests that affective learning about places is partially based on general associative mechanisms, such as classical conditioning (for reviews, see Pearce & Bouton, 2001; Dickinson & Mackintosh, 1978) and evaluative conditioning (for reviews, see De Houwer et al., 2001; Hofmann et al., 2010), by enabling the acquisition of new knowledge over repeated experiences of simultaneous co-occurrence of a neutral and an affective stimulus. However, Experiment 1 showed that basic affective mechanisms are not the only way through which we can learn about the inanimate environment. In fact, participants learned about places when image-statement pairs were presented only once (condition in which *CC* and *EC* are less likely to occur). In this case, affective associations with images of places were created only when the information was relevant (note that, however, this experiment might have been underpowered). In line with this, Experiments 2 and 4 indicated that affective learning about places and houses is boosted by the relevance of the statement to the image. Since little evidence exists relative to how affective associations with non-human targets are triggered, the exact mechanisms underpinning this process remain unclear and warrant further research.

Lastly, in Experiment 2 and Experiment 3, we explored whether varying the type of behavioral information from which perceivers could extract affective value would affect how individuals can learn about others. Unfortunately, this manipulation did not lead to clear findings (and it did not qualify the main results concerning affective learning of associations with people and places). Future research is needed to better clarify this issue.

In sum, our studies demonstrate that affective learning of both people and non-human targets partially depend on general affect-based associative mechanisms (in the case of multiple repetitions of the affective information). However, this general valence-based effect seems to be more powerful in triggering affective associations with non-human than with human targets, possibly indicating that the processes underlying the affective learning about people go beyond general affect-based mechanisms.

## Data availability

All data are available on Open Science Framework: https://osf.io/aqhxz/.

[2] In interpreting our findings, it should be taken in consideration that we used the same evaluation scale (goodness) for the judgment of both human and non-human targets. This choice was in line with our previous study (Falvello et al., 2015) and allowed us to keep the task perfectly identical for the two categories of stimuli (human and inanimate). However, goodness might have different meanings when referred to people and inanimate targets, possibly emphasizing differences between the two categories.

## Credit authorship contribution statement

**Chiara Ferrari**:Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing, Validation.**DongWon Oh**:Investigation, Formal analysis, Writing - original draft, Writing - review & editing.**Brandon P. Labbree**:Investigation, Writing - review & editing.**Alexander Todorov**:Methodology, Formal analysis, Writing - original draft, Writing - review & editing.

## Declaration of competing interest

None.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.actpsy.2020.103011.

## References


Bliss-Moreau, E., Barrett, L. F., & Wright, C. I. (2008). Individual differences in learning the affective value of others under minimal conditions. *Emotion, 8*, 479–493.

Bruce, V., Doyle, T., Dench, N., & Burton, M. (1991). Remembering facial configurations. *Cognition, 38*(2), 109–144.

Carlston, D. E., & Skowronski, J. J. (2005). Linking versus thinking: Evidence for the different associative and attributional bases of spontaneous trait transference and spontaneous trait inference. *Journal of Personality and Social Psychology, 89*, 884–898.

Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Wunder, J., & De Rosario, H. (2015). Basic functions for power analysis. R package versions. 1.2-1. Retrieved from https://cran.rproject.org/web/packages/pwr/pwr.pdf.

De Houwer, J., Thomas, S., & Baeyens, F. (2001). Association learning of likes and dislikes: A review of 25 years of research on human evaluative conditioning. *Psychological Bulletin, 127*(6), 853–869.

Dickinson, A., & Mackintosh, N. J. (1978). Classical conditioning in animals. *Annual Review of Psychology, 29*, 587–612.

Falvello, V., Vinson, M., Ferrari, C., & Todorov, A. (2015). The robustness of learning about the trustworthiness of other people. *Social Cognition, 33*, 368–386.

Gawronski, B., & Bodenhausen, G. V. (2006). Associative and propositional processes in evaluation: An integrative review of implicit and explicit attitude change. *Psychological Bulletin, 132*, 692–731.

Ghuman, A. S., & Bar, M. (2006). The influence of nonremembered affective associations on preference. *Emotion, 6*(2), 215–223.

Hofmann, W., De Houwer, J., Perugini, M., Baeyens, F., & Crombez, G. (2010). Evaluative conditioning in humans: A meta-analysis. *Psychological Bulletin, 136*(3), 390–421.

Konkle, T., Brady, T. F., Alvarez, G. A., & Oliva, A. (2010). Scene memory is more detailed than you think: The role of categories in visual long-term memory. *Psychological Science, 21*, 1551–1556.

Landy, J. F., Piazza, J., & Goodwin, G. P. (2018). Morality traits still dominate in forming impressions of others. *Proceedings of the National Academy of Sciences, 115*(25), E5636.

Levey, A. B., & Martin, I. (1975). Classical conditioning of human “evaluative” responses. *Behaviour Research and Therapy, 13*(4), 221–226.

Mattarozzi, K., Colonnello, V., Russo, P. M., & Todorov, A. (2018). Person information facilitates memory for face identity. *Psychological Research,* 1–8.

Maurer, D., Le Grand, R., & Mondloch, C. J. (2002). The many faces of configural processing. *Trends in Cognitive Sciences, 6*(6), 255–260.

McCarthy, R. J., Wells, B. M., Skowronski, J. J., & Carlston, D. E. (2018). Multiple behavior descriptions affect the acquisitions of STI and STT. *Psychological Reports, 121*(4), 615–634.

Olson, M. A., & Fazio, R. H. (2001). Implicit attitude formation through classical conditioning. *Psychological Science, 12*(5), 413–417.

Orghian, D., Garcia-Marques, L., Uleman, J. S., & Heinke, D. (2015). A connectionist model of spontaneous trait inference and spontaneous trait transference: Do they have the same underlying processes? *Social Cognition, 33*(1), 20–66.

Pavlov, I. P. (1927). *Conditioned reflexes: An investigation of the physiological activity of the cerebral cortex (G. V. Anrep, Trans.).* Oxford, UK: Oxford University Press.

Pearce, J. M., & Bouton, M. E. (2001). Theories of associative learning in animals. *Annual Review of Psychology, 52*, 111–139.

Phillips, P. J., Moon, H., Rizvi, S. A., & Rauss, P. J. (1999). *The FERET evaluation methodology for face-recognition Algorithms 1 introduction.* Technical reportNational Institute of Standars and Technology.

Pitcher, D., Walsh, V., Yovel, G., & Duchaine, B. (2007). TMS evidence for the involvement of the right occipital face area in early face processing. *Current Biology, 17*(18), 1568–1573.

Schneid, E. D., Carlston, D. E., & Skowronski, J. J. (2015). Spontaneous evaluative inferences and their relationship to spontaneous trait inferences. *Journal of Personality and Social Psychology, 108*(5), 681–696.

Shigihara, Y., & Zeki, S. (2014). Parallel processing of face and house stimuli by V1 and specialized visual areas: A magnetoencephalographic (MEG) study. *Frontiers in Human Neuroscience, 8*, 901.

Skowronski, J. J., Carlston, D. E., Mae, L., & Crawford, M. T. (1998). Spontaneous trait transference: Communicators take on the qualities they describe in others. *Journal of Personality and Social Psychology, 74*(4), 681–696.

Summerfield, C., Egner, T., Mangels, J., & Hirsch, J. (2005). Mistaking a house for a face: Neural correlates of misperception in healthy humans. *Cerebral Cortex, 16*(4), 500–508.

Todorov, A., & Olson, I. R. (2008). Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged. *Social Cognitive and Affective Neuroscience, 3*, 195–203.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition, 27*, 813–833.

Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actors' faces: Evidence from a false recognition paradigm. *Journal of Personality and Social Psychology, 83*, 1051–1065.

Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actors' faces. *Journal of Experimental Social Psychology, 39*, 549–562.

Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology, 87*, 482–493.

Uhlmann, E. L., Pizarro, D. A., & Diermeier, D. (2015). A person-centered approach to moral judgment. *Perspectives on Psychological Science, 10*, 72–81.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100 ms exposure to a face. *Psychological Science, 17*, 592–598.

Wojciszke, B., Bazinska, R., & Jaworski, M. (1998). On the dominance of moral categories in impression formation. *Personality and Social Psychology Bulletin, 24*, 1251–1263.