UNIVERSITY of CALIFORNIA PRESS | Collabra: Psychology

Rougier, M., De Houwer, J., Richetin, J., Hughes, S., & Perugini, M. (2023). From Halo to Conditioning and Back Again: Exploring the Links Between Impression Formation and Learning. *Collabra: Psychology*, *9*(1). https://doi.org/10.1525/collabra.84560

Social Psychology

From Halo to Conditioning and Back Again: Exploring the Links Between Impression Formation and Learning

Marine Rougier[1][a], Jan De Houwer[1], Juliette Richetin[2], Sean Hughes[1], Marco Perugini[2]

[1] Department of Experimental Clinical and Health Psychology, Ghent University, Ghent, Belgium, [2] Department of Psychology, University of Milano-Bicocca, Milano, Italy

Keywords: Impression formation, Evaluative conditioning, Attribute conditioning, Halo effect, Pairing-based halo

https://doi.org/10.1525/collabra.84560

Collabra: Psychology

Vol. 9, Issue 1, 2023

Impression formation effects – such as the halo effect – and learning effects – such as evaluative or attribute conditioning effects – are often seen as separate classes of phenomena. In a recent conceptual paper, De Houwer et al. (2019) suggested that both may actually qualify as instances of feature transformation, where a source feature (e.g., attractiveness of a face; valence of an unconditioned stimulus; US) influences judgements about a target feature (e.g., social competence of a person; valence of a conditioned stimulus; CS). In halo effects, the source and target features typically differ (e.g., a person with an attractive face is judged as more socially competent) but belong to the same object. In evaluative conditioning, source and target features are the same (e.g., a neutral CS is judged as more positive after being paired with a positive US) but belong to different objects. In this paper, we highlight a phenomenon at the crossroads of the two previous effects: feature transformation where source and target features are different (as in halo studies) and belong to different objects that are paired together (as in evaluative conditioning studies). Across six pre-registered experiments (n = 1050), we obtained evidence for this phenomenon in the context of person perception (i.e., attractiveness halo) and food perception (i.e., health halo). We also show that this type of feature transformation is influenced by several known moderators of halo and conditioning effects (beliefs about traits relationship, memory of pairings, and salience of the source feature).

Humans often have little difficulty making judgements about features of the stimuli that surround them, even if those features are unknown or ambiguous. Research on impression formation revealed that judgements about unknown or ambiguous stimulus features are often biased by other information about the stimulus. Studies on the halo effect, for instance, show that physically attractive persons are judged as being more socially competent than less attractive persons (e.g., Dion et al., 1972; Eagly et al., 1991). Similar phenomena have been observed in evaluative learning research. For instance, evaluative conditioning (EC) studies showed that judgements about the valence of a neutral conditioned stimulus (CS) are influenced by the valence of an unconditioned stimulus (US) with which the CS is paired (for reviews, see De Houwer et al., 2001; Hofmann et al., 2010).

Recently, De Houwer et al. (2019) proposed a conceptual framework that unifies impression formation and learning phenomena under the "feature transformation" umbrella. This proposal led us to examine a new phenomenon that combines elements from both types of phenomena. We not only provide the first empirical evidence for this new type of feature transformation effect but also examine several potential moderators. Our work illustrates the possible cross-fertilization of learning and impression formation research. Before discussing our studies in more detail, we first provide information about halo effects, conditioning effects, and how they relate to each other from the perspective of feature transformation.

Halo Effects

The halo effect refers to a situation where a positive characteristic of a stimulus influences how the stimulus is perceived on other dimensions for which no information is available (Forgas & Laham, 2016). Although the halo effect is sometimes described as a broad phenomenon in the literature (e.g., Nisbett & Wilson, 1977), our work focuses on two specific cases of halo effects (attractiveness

a Correspondence should be addressed to Marine Rougier, Department of Experimental Clinical and Health Psychology, Ghent University, H.Dunantlaan 2, 9000 Ghent, Belgium. Email: Marine.Rougier@UGent.be.

and health halo effects). A typical example of halo is the effect of physical attractiveness on personality perception, initially called the "what is beautiful is good" effect (Dion et al., 1972): Attractive people are usually judged as having more desirable traits (e.g., intelligence) and, as a result, receive a more positive treatment than unattractive people (e.g., Efran, 1974; Landy & Sigall, 1974; for meta-analyses, see Eagly et al., 1991; Langlois et al., 2000).

It has been argued that the attractiveness halo effect provides a window into implicit personality theories, that is, the beliefs on how traits should co-vary among individuals (e.g., friendly people should also be trustworthy; Rosenberg et al., 1968; see also Forgas & Laham, 2016). The halo effect has also been related to stereotyping, where stereotypes can be defined as the relationship between a group (e.g., attractive people) and other attributes (e.g., sociable), leading to inferences about (stereotyping of) the personality of a member of that group (e.g., this person is sociable because s/he belongs to the attractive people group; Ashmore, 1981). However, the attractiveness halo effect is not evident for all personality dimensions. For instance, in their meta-analysis, Eagly et al. (1991) found that the influence of physical attractiveness is larger for traits related to social competence compared to adjustment, potency, or intellectual competence, which in turn are influenced more strongly than traits relating to concern for others and integrity. Moreover, and in contrast to the "what is beautiful is good" idea, attractiveness can sometimes lead to negative judgements on certain traits. For instance, attractive people are judged as being more vain, less competent in parenting, and less likely to be faithful to their partner (Bassili, 1981; Dion et al., 1972; Han & Laurent, 2023; see also Sigall & Ostrove, 1975). This selectivity in how certain traits influence the assumptions that people make about other traits (what we call 'trait selectivity') is thought to reflect the structure of personality theories or the stereotypes that people have about attractive versus unattractive people (e.g., Dermer & Thiel, 1975; Eagly et al., 1991).

Although initially studied in social judgement research, the halo effect is also evident in the context of food products. For instance, take the "health halo" effect: labels such as "low in fat", "rich in fibre", "without cholesterol", or even "organic" often lead people to underestimate the calorie content of the food products to which those labels are applied, ultimately leading to over-consumption (Andrews et al., 1998; Forwood et al., 2013; Lee et al., 2013; Wansink & Chandon, 2006). Trait selectivity is evident here too: there is generally a larger label effect (e.g., organic vs. regular) on health-related traits compared to hedonic traits (e.g., tasty), on which there can be a reversed effect (i.e., lower ratings for organic food; Lee et al., 2013; Orquin & Scholderer, 2015; Westcombe & Wardle, 1997; but see Ebneter et al., 2013).

Evaluative Conditioning

Elsewhere, classical conditioning effects have long been at the core of research on learning. Although there are many different types of conditioning effects, they all have in common the fact that behaviour changes as the result of pairing stimuli (i.e., presenting stimuli together in space and time; see De Houwer & Hughes, 2020, for a review). In Evaluative Conditioning (EC), for instance, the evaluation of an initially neutral stimulus (CS) changes due to repeated pairings with a liked or disliked stimulus (US; De Houwer, 2007). Whereas EC refers to changes in evaluation, attribute conditioning (AC) refers to changes in judgements about specific attributes (e.g., athleticism; Förderer & Unkelbach, 2011, 2015). For instance, repeated pairings of neutral individuals (CSs) with athletic versus nonathletic people (USs) influence the extent to which the CSs are viewed as being athletic (Förderer & Unkelbach, 2011). Something similar is also true for fear conditioning, disgust conditioning, and other conditioning effects (i.e., a change in behaviour along a particular dimension as a result of stimulus pairings).

Historically, conditioning effects have typically been explained according to one of two perspectives. Some have argued that they result from the formation and activation of an association between the representations of the CS and the US (e.g., Rescorla & Wagner, 1972; see also Gast & Rothermund, 2011) while others argue that they result from the formation and activation of propositions about relations in the environment (e.g., the proposition "the CS and US go together"; De Houwer, 2009, 2018).

Conditioning effects are known to be sensitive to many moderators (see Hofmann et al., 2010, for a review). For instance, EC and AC effects are generally larger when participants correctly remember which CS and US were paired together during the pairing procedure (e.g., Förderer & Unkelbach, 2014; Stahl & Unkelbach, 2009). Moreover, EC and AC effects strongly depend on which stimulus attributes are salient during the pairing procedure. EC effects seem to emerge only when participants focus on valence during the learning phase (e.g., when they judge the likability of persons but not their geographic origin; Gast & Rothermund, 2011; see also Vanaelst et al., 2017). Similarly, when USs possess multiple attributes (e.g., athletic vs. intelligent) in AC studies, which of these attributes are conditioned depends on which are made salient during the experiment (e.g., Förderer & Unkelbach, 2014; see also Olson et al., 2009).

The Feature Transformation Framework

Halo and conditioning effects emerged from two different parts of psychological science (i.e., social psychology and learning psychology, respectively). They are typically described using different terms, have been explained on the basis of different cognitive theories, and research on these phenomena has focused on different types of moderators. Nevertheless, as De Houwer et al. (2019) pointed out, they both relate to how assumptions about unknown or ambiguous features of stimuli are shaped by known aspects of the environment. Based on this idea, De Houwer et al. (2019) recently proposed a conceptual framework that can be used to describe both halo and EC effects. This framework introduces four generic, theory-free core concepts that can be used to describe and classify a wide range of phenomena, including halo and conditioning effects. These abstract concepts refer to a source feature, target feature, source object, and target object.

According to this framework, a target feature is a feature of an object about which people make assumptions. It belongs to the target object. The source feature is a feature that influences the assumptions people make about the target feature. It belongs to the source object. Importantly, “features” can relate to various aspects of stimuli (e.g., physical, psychological, and behavioural aspects), and “objects” can relate to various types of stimuli (e.g., faces, persons, animals, products). In most cases, the source feature is the one that we vary as the independent variable (e.g., by presenting faces varying on their physical attractiveness), and the target feature is the one we assess as a dependent variable (e.g., how much each face is perceived as socially competent). When the source feature influences assumptions about the target feature (e.g., the more attractive the person, the more socially competent s/he is perceived), feature transformation is said to have occurred, independent of the direction of the effect. A crucial aspect of the framework is that it does not make claims about the mental processes that influence the source’s impact on the target (e.g., intermediary mechanisms). The term feature “transformation” is simply used to describe an *effect* (i.e., the impact of the manipulated source feature on a given target feature) independent of what drives this effect at the mental level.

The feature transformation framework not only clarifies what halo and conditioning effects have in common but also how they differ. Take, for instance, the attractiveness halo effect. Here the source and target features differ whereas the source and target objects are the same: a person has a source feature (e.g., scores high, medium, or low on attractiveness) and is judged on another (target) feature (e.g., social competence, Dion et al., 1972). The same is true for health halo effects (e.g., the food product that is said to have a high, medium, or low-fat content, is judged on how healthy it is; Ebneter et al., 2013). However, EC and AC effects involve a different situation, one where the source and target features are identical, but the source and target objects differ. For instance, in EC, valence is both the manipulated (source) feature and the measured (target) feature, but the former is part of the US, and the latter is part of the CS. The same is true for AC except that the nature of the (source and target) feature is different (e.g., an attribute such as athleticism rather than valence). Finally, in both EC and AC studies, the source object (US) and target object (CS) are related by pairing them (i.e., by presenting them together in space and time).

New Ideas for Research

The feature transformation framework reveals not only new ways of relating seemingly unrelated phenomena such as halo and conditioning effects but also new phenomena that have yet to be identified and studied. In this paper, we introduce one such effect: the pairing-based halo effect.

This effect involves feature transformation wherein (a) the source and target features differ, (b) the source and target objects differ, and (c) the source and target objects are related via pairings. This – in effect – involves a combination of the halo effect and conditioning effects. Imagine, for instance, that we pair the picture of a highly attractive person with the picture of a moderately attractive person. As a result of the pairings, the moderately attractive person is judged to be more socially competent afterwards. Such an effect resembles halo effects in that the source feature (attractiveness) and target feature (social competence) are different. It resembles conditioning effects in that the source object (attractive person) and target object (moderately attractive person) are different but paired together. The current paper provides the first empirical demonstration of such an effect and examines several of its moderators. Because the feature transformation framework is situated at the descriptive level (not at the process level), our goal is not to investigate mental processes underlying the pairing-based halo effect. Potential avenues for mental processes will be covered in the General Discussion.

Of note, the pairing-based halo effect bears a certain resemblance to a halo effect known as the “radiating beauty” effect. Research shows that a woman’s attractiveness influences perceptions of her male partner: the more attractive the woman, the more positively her male partner is judged on a series of target features (Kocoglu & Mithani, 2020; Sigall & Landy, 1973; see also Walther et al., 2008). Both effects rely on different source and target features (attractiveness and self-confidence) and different source and target objects (men and women). However, in the radiating beauty effect, the source and target objects are related via instructions (e.g., they are said to be partners), whereas in the pairing-based halo effect, they are related via mere spatiotemporal contiguity (i.e., two people are presented together in space and time). Hence, only the pairing-based halo effect involves the pairing procedure that is typical of conditioning studies.

Given that it contains elements of both halo and conditioning effects, the pairing-based halo effect is an ideal example of how the feature transformation framework can create a cross-fertilization between distinct research areas. Hence, we tested whether this hybrid feature transformation effect is sensitive to those factors that moderate halo and conditioning effects. To illustrate this, take the idea of ‘trait selectivity’ in halo effects which we mentioned previously (i.e., that certain inferences [social competence] follow when we learn about a trait [attractiveness] whereas others do not [integrity]). Will a similar trait selectivity be observed in the pairing-based halo effect? Or, will the attractiveness of the source object functions only as a positive stimulus whose valence is transferred to the target object that it is paired with (as in EC)? Similarly, will the pairing-based halo effect be larger when participants have a correct (vs. incorrect) memory of the source and target object pairs and when the source feature (e.g., attractiveness) is made salient, as is the case for conditioning effects?

Examining such questions enriches the literature on impression formation by highlighting and underlying stimulus pairings as a pathway for the formation of first impressions. Likewise, it expands the literature on conditioning by stressing that the effects of stimulus pairings can go beyond the feature of the US that is being manipulated. This new line of research could also lead to new insight into the processes that play a role in impression formation and conditioning. For instance, whereas pre-existing conceptual beliefs such as implicit personality theories and stereotypes are thought to be central in impression formation, their role has been largely unexplored in the conditioning literature. Relating the two phenomena raise the possibility that pairing effects are also based on pre-existing conceptual beliefs. We will first present our studies and findings and then return to these theoretical considerations in the General Discussion.

Current Studies

As the first step in this new line of research, we set out to replicate classic attractiveness and health halo effects as well as their respective trait selectivity in Studies 1 and 5. Based on the existing literature, we relied on target traits that are known to be more or less "halo relevant" (i.e., low vs. medium vs. high relevance), which indicates to what extent a given trait typically produced a halo effect in past research (i.e., whether it produced a small vs. intermediary vs. large effect). For instance, in the attractiveness halo effect, traits related to sociability typically produced a larger halo effect (i.e., a larger difference of rating when faces are high vs. low on attractiveness) compared to integrity (Eagly et al., 1991). This means that traits related to sociability are more "halo relevant" than those related to integrity. Importantly, we also verified that the observed trait selectivities were not explained by a mere valence effect – what has been argued but never addressed directly in the halo literature. To test this alternative hypothesis, we assessed the extent to which each target trait was generally perceived to be positive or negative, what we refer to as "trait valence" (see Pilot study in Study 1 and Study 5), and tested whether the observed halo effects were moderated by trait halo relevance or by trait valence.

Then, Studies 2-4 (attractiveness halo) and Study 6 (health halo) sought evidence for the pairing-based halo effect. In these studies, participants first encountered a pairing procedure similar to that used in conditioning studies. Pairs of faces or products were repeatedly presented onscreen, with one (target) object being relatively neutral in terms of its attractiveness or calorie content and the other (the source object) varying on this dimension (i.e., low vs. medium vs. high; the source feature). Then, we asked participants to judge the target objects (e.g., the faces neutral on attractiveness) on a series of dimensions (target features) using scales from 0 (e.g., "unlikely to achieve career success") to 5 (e.g., "likely to achieve career success"). We also aimed to investigate three potential moderators of the pairing-based halo effect.

A first moderator was the trait selectivity: we examined if the trait selectivity observed in the two classic halo effects also emerges when a pairings-based procedure is used. Whereas in Studies 1 and 5 we relied on categories of halo relevance (i.e., low vs. medium vs. high relevance) determined based on previous literature, in the pairing-based studies we relied on the results obtained in Studies 1 and 5. Specifically, we calculated for each target feature an average halo effect score. For instance, in the pairing studies on attractiveness, we used the average ratings for each trait (e.g., sociability) when the target face was paired with a high attractive source face vs. when it was paired with a low attractive source face and computed a score of difference between the two (e.g., trait relevance for sociability = rating when the source face was high – rating when the source face was low). We used this score per trait as an index of "trait halo relevance" because it codes for the extent to which a given target feature (e.g., social competence) is influenced by (i.e., more or less relevant for) the source feature (e.g., attractiveness) in a classic halo procedure. One advantage of relying on continuous scores that are data-driven (because derived from our own datasets) is that the scores are derived from our own set of faces or products that are used as source objects in the pairing-based studies.

In the studies that followed (Studies 2, 3, 4, and 6), we tested whether the pairing-based halo effect increased as a function of the trait halo relevance (e.g., the higher the trait halo relevance, the higher the pairing-based halo effect). Such a result would demonstrate that the pairing-based halo effect is "relevance-based" (i.e., similar to the standard halo effect in terms of trait selectivity). Alternatively, other types of selectivity could be observed: given that the source faces possess multiple features, it could be the case that other features are used by the participants, leading to a different trait selectivity on the target face (e.g., Förderer & Unkelbach, 2014). One likely alternative is that pairing-based effects are merely "valence-based", in line with what would be observed in the case of EC effects – where the manipulated source feature is the valence of the source object. As we did for Studies 1 and 5 in the case of typical halo effects, we systematically tested whether the observed pairing-based effects were moderated by trait halo relevance or trait valence.

The two other moderators that we investigated are the ones that typically impact conditioning effects: (a) memory of source and target object pairings and (b) salience of the source feature. We expected pairing-based halo effects to be larger when participants correctly (vs. incorrectly) remembered the stimuli that were paired and when the source feature was made salient (vs. non-salient). Finally, we tested if these two moderators also influenced the type of trait selectivity that takes place (i.e., relevance-based vs. valence-based).

Transparency, Openness, Analytical Strategy, and Diversity of Samples

We pre-registered all our studies on Open Science Framework or AsPredicted. Pre-registrations include a priori theoretical reasoning, hypotheses, power estimations, procedures, and statistical analyses. Data were analysed using RStudio, version 1.4.1106 (RStudio Team, 2021). We report for each study how we determined the sample size, any data exclusions, manipulations, and measures. The main document highlights major deviation from the initial pre-registrations, while minor deviations are presented in a separate file. The pre-registration files, deviations from the pre-registrations document, materials (including lists of stimuli and JsPsych or Inquisit code), data, and analytic (R) scripts for all experiments have been made publicly available at https://osf.io/5dftn/?view_only=fc2c3f2e2e0e402d93d7d2b81a57a284. Studies on the attractiveness halo effect received approval (number 2021/39) from the ethical committee of the Faculty of Psychology and Educational Sciences at Ghent University, and studies on the health halo effect received approval (number RM-2021-395) from the Faculty of Psychology at University of Milan-Bicocca.

A major deviation from the pre-registrations involves the coding of trait halo relevance in the pairing-based halo studies (Studies 2-4 and 6). We initially planned to use the categorical variable of trait halo relevance used in the classic halo studies (Studies 1 and 5). Instead, as explained before, we decided to rely on continuous scores of trait halo relevance obtained in Studies 1 and 5 (i.e., the extent to which each trait produced a halo effect). This strategy enabled us to use a finer-grained measure of a trait halo relevance at the trait level instead of the category (of trait) level. Importantly, having a continuous measure was also more adequate for testing whether the obtained pairing-based effects were relevance-based (i.e., varied as a function of a trait halo relevance) or valenced-based (i.e., varied as a function of a trait valence). Indeed, this enabled to clearly oppose two competing hypotheses with similar chances for both effects to emerge. Both variables were continuous and situated at the trait level, allowing us to disentangle both effects beyond potential collinearity (Irwin & McClelland, 2003). A continuous measure also decreases Type I and II errors compared to a categorical variable (Irwin & McClelland, 2003; see also McClelland et al., 2015).

This change, however, implied adapting our statistical strategy. Instead of conducting the pre-registered OLS regressions, we opted for mixed-model analyses (Judd et al., 2017). This type of analysis made it possible to use continuous within-participant variables (here, trait halo relevance and trait valence) because it models the non-independence of the residuals for participants (by estimating random slopes). In contrast, OLS regressions would have required composite scores and thus a loss of information and statistical power. In addition, mixed-model analyses enable the use of multiple random factors (i.e., participants and traits) instead of one (e.g., only participants) as in more traditional analyses of variance. Mixed-models thus enhance the robustness and the generalisability of the findings because they allow generalising the results to other participants and traits simultaneously (Judd et al., 2012; Westfall et al., 2014). We used mixed-model analyses to test all our hypotheses, except for the attractiveness or fat content check for which we used an OLS regression because this analysis could only have participants and not traits as a random factor (i.e., it involved ratings on one or two traits only). For the sake of transparency, we report the pre-registered analyses as Supplementary Materials, using the categorical variable of trait relevance and using both mixed-models and OLS regression. The use of these various analytical strategies does not impact our main conclusions.

All our experiments were conducted online, either via the crowdsourcing platform Prolific Academic (https://prolific.ac/) or via the network of University of Milan-Bicocca University. Samples included native English-speaking participants (Studies 1-4) or Italian participants (Studies 5-6). We had no reasons to predict different effects as a function of the characteristics of our samples (e.g., gender, socioeconomic variables, etc.). It is likely that our samples of participants were quite heterogeneous in the sense that participants potentially greatly vary on gender, religion, and regarding their belonging to ethnic and racial groups, social class, and so forth (Gleibs, 2017; Peer et al., 2017). However, participants from Studies 1-4 varied only slightly on age as we selected participants aged between 18 and 30, based on the published literature on the halo effect. As such, future work should test if the present set of results can be verified on an older population.

Studies 1-4: Attractiveness Halo Effects

Study 1

Study 1 set out to replicate the classic attractiveness halo effect. We mimicked the procedure used in Dion et al. (1972). More specifically, we assessed the impact of facial attractiveness as a source feature on 42 target features that are known to be influenced by attractiveness to a greater or lesser extent (Bassili, 1981; Eagly et al., 1991). In this study, the source and target faces[1] were always the same (i.e., people who varied along the source feature of attractiveness were rated on a series of target features). Based on the results of earlier halo studies, we expected that the more attractive the face, the more positive the ratings, with the exception of vanity-related traits. We also expected the halo effect to vary as a function of the target feature being assessed (i.e., to observe trait selectivity). That is, we expected certain trait judgements to be influenced by attractiveness more than others in ways that are consistent with the attractiveness stereotype. For instance, based on the meta-analysis of Eagly et al. (1991), we expected that traits relating to social competence should be more influenced than traits relating to integrity. We also tested whether trait selectivity was still present even after controlling for the valence of the traits. If the classic halo effect is based on the attractiveness stereotype, then attractiveness should lead to both positive (e.g., more socially competent) and negative (e.g., more vain) ratings as a function of the feature being assessed. If the effect merely relies on valence, then the impact of attractiveness on trait ratings should be a func-

1 For the sake of simplicity, when referring to specific studies, we employ the terms “source/target face” (Studies 1-4) or “source/target product” (Studies 5-6) instead of “source/target object”.

tion only of the valence of the trait (i.e., higher ratings for more positive traits).

Method

Participants and design. To estimate the required sample size for sufficient power (80%), we relied on the mean weighted effect size of the average halo effect (*ds* = 0.58) found in Eagly et al.'s (1991) meta-analysis. We recruited 100 participants (M_{age} = 25.24, SD_{age} = 3.64, 56 women, 43 men, and 1 participant responding "other"), giving us 80% likelihood of detecting a minimum effect of *dz* = 0.25. Participants took part in exchange for £1.50 and were recruited via the Prolific Academic platform (www.prolific.co). Participants spoke English as their first language, did not take part in any other study from our lab, and had an approval rate of at least 90% (this last criterion leads to improved data quality; Peer et al., 2014). In line with studies on the attractiveness halo effect that mainly rely on college undergraduates (Bassili, 1981; Dermer & Thiel, 1975; Dion et al., 1972), we recruited participants aged between 18 and 30. A 3 (*Attractiveness*: low vs. medium vs. high) x 3 (*Trait Relevance*: low vs. medium vs. high) x continuous (*Trait Valence*: from -3 to +3) x 2 (*Target Gender*: male vs. female) design was used with the final control variable manipulated between-participants.

Materials. We selected coloured photographs of six male and six female faces from the 10k US Adult Faces Database (Bainbridge et al., 2013), that is, two faces per level of physical attractiveness and per gender. Faces were selected to vary on attractiveness but to not significantly differ on a series of other trait dimensions (*see below*; all scales ranged from 1 = *not at all* to 9 = *extremely*). We compared the faces using two orthogonal contrast codes: a linear contrast C1 opposing low with high attractiveness conditions (low = -1/2, medium = 0, high = 1/2) and a quadratic C2 opposing low and high conditions taken together with the medium condition (low = -1/3, medium = 2/3, high = -1/3). Low and high attractive faces significantly differed on attractiveness, $t(9) = 3.04$, $p = .014$, $d = 1.01$, 95% CI [0.20; 1.79], whereas the medium faces did not significantly differ from low and high attractiveness faces taken together, $t(9) = 0.59$, $p = .57$, $d = 0.07$, 95% CI [-0.58; 0.73] (M_{low} = 4.57, SD_{low} = 0.43, M_{med} = 5.35, SD_{med} = 1.64, M_{high} = 6.92, SD_{high} = 0.83). The faces did not significantly differ (all *ps* < .22) on emotional intensity (M_{low} = 3.35, SD_{low} = 0.49, M_{med} = 3.14, SD_{med} = 0.69, M_{high} = 2.85, SD_{high} = 0.99), image quality (M_{low} = 3.91, SD_{low} = 0.04, M_{med} = 3.62, SD_{med} = 0.48, M_{high} = 3.94, SD_{high} = 0.40), memorability (M_{low} = 3.08, SD_{low} = 0.31, M_{med} = 3.06, SD_{med} = 0.35, M_{high} = 3.44, SD_{high} = 0.45), and quantity of teeth visible (M_{low} = 0.75, SD_{low} = 0.50, M_{med} = 0.75, SD_{med} = 0.96, M_{high} = 1.00, SD_{high} = 1.15). Faces were all white, unknown (i.e., not celebrities), and most of them belonged to the 20-30 years old category.

Personality traits and outcomes were selected from Eagly et al. (1991) and Bassili (1981) and varied along seven dimensions: social competence (sociable, fun-loving, likable, popular), vanity/materialistic orientation (elitist, snobbish, shallow, humble, materialistic, pompous, prudish, boastful, vain), adjustment (normal, well-adjusted, satisfied, happy, confident, having a positive self-regard, mature, healthy), potency (strong, self-assertive, dominant, leader), intellectual competence (intelligent, skillful, rational, scientific, ambitious, hard-working, likely to receive good grades, likely to achieve career success), concerns for others (sensitive, empathic, compassionate, generous, modest, egoistic), and integrity (trustworthy, honest, likely to be faithful to the spouse). Based on Eagly et al. (1991) and Bassili's (1981) results (i.e., observed halo effect sizes), we classified traits into three categories of relevance regarding the source feature of attractiveness: 1) traits relating to social competence and vanity were "highly relevant" (i.e., they should produce a large halo effect), 2) traits relating to adjustment, potency, and intellectual competence were "medium relevant" (i.e., they should produce an intermediary halo effect), and 3) traits relating to integrity and concerns for others were "low relevant" (i.e., they should produce a small effect or no effect at all).

In a pilot study (N = 40, M_{age} = 40.13, SD_{age} = 17.74, 20 women and 20 men), we collected ratings on trait valence (i.e., to what extent a trait was positive or negative) using a scale that ranged from -3 (*extremely negative*) to +3 (*extremely positive*), self-relevance (i.e., to what extent a trait is consequential for the trait holder), other-relevance (i.e., to what extent a trait is consequential for the individuals living nearby the trait holder; both of which ranged from 1 [*low consequences*] to 7 [*high consequences*]), and face-readability (i.e., to what extent is it easy to infer a personality trait on the basis of someone's face; from 1 [*not easy at all*] to 7 [*extremely easy*]). We obtained these trait ratings to exclude their potential effects.

Procedure. The experiment was programmed using jsPsych (de Leeuw, 2015). After giving their consent, participants received instructions similar to typical attractiveness halo studies (e.g., Dion et al., 1972). Specifically, they were informed that we were interested in the accuracy of person perception and that our aim was to determine the extent to which impressions about other people are generally accurate. Participants were told that their accuracy in person perception would be compared with other groups who had been trained in various interpersonal perception techniques (students in clinical psychology and professional clinical psychologists). We told them that certain individuals without training might be as accurate as some professionals. Finally, participants were told that the photographs they would encounter were part of a group of college students currently enrolled in a longitudinal study of personality development, so it would be possible to assess their judgement accuracy by comparing judgements with people's real behaviour.

Following these instructions, participants encountered three photographs presented one at a time in random order. The photographs were randomly selected from the pool of faces so that, for a given participant, all of them showed male or female faces, and one picture was included for each of the three levels of attractiveness. Participants evaluated each face on the 42 different personality traits and outcomes (all scales ranged from 0 = *not at all* to 5 = *totally*). Faces were displayed at the top of the screen, with the rating scales below. Once they had completed all ratings for one face, they moved on to the next face. The order in which the personality traits and outcomes were presented was randomised for each face and each participant separately. Participants were encouraged to answer as honestly and as spontaneously as possible. At the end of the rating phase, participants were asked to rate the three faces on their attractiveness (from -3 = *extremely unattractive* to +3 = *extremely attractive*). Again, faces were displayed at the top of the screen, with the attractiveness scale below. They then provided demographic information (age, gender, English fluency), left optional comments about the study, and were thanked and debriefed.

Results

We first reversed the ratings for negatively valenced traits, except for vanity-related traits, then calculated the average across all 42 ratings. We expected the average rating to be higher for attractive than for non-attractive faces. In addition, we predicted the effect of attractiveness to increase as a function of the trait relevance. Reliabilities within each trait dimension were relatively high (from α = .71, 95% CI [.62; .81], for social competence to α = .91, 95% CI [.89; .94], for intellectual competence). The variables of attractiveness and trait relevance were both coded via two orthogonal contrast codes (linear contrast C1: low = -1/2, medium = 0, high = 1/2; quadratic contrast C2: low = -1/3, medium = 2/3, high = -1/3). We focused on the linear contrast code opposing the low with the high for attractiveness and dimension relevance (i.e., low vs. high attractive faces; low vs. high relevant traits). The results for the quadratic contrasts are presented in the Supplementary Materials. The control factors of participant gender, trait self-other relevance, and face readability did not significantly moderate the halo effect or the trait selectivity in any of our studies – with exceptions mentioned in the main text. Hence we removed them from the presented analyses. We report the results for the attractiveness check, the halo effect, and its relevance- vs. valence-based trait selectivity. Mixed-model analyses were performed using the lmerTest package version 3.1-0 (Kuznetsova et al., 2017).

Attractiveness ratings (OLS regression). In line with our categorisation based on Bainbridge et al.'s (2013) database, participants judged high attractive faces as more attractive than low attractive faces, $t(99) = 13.09$, $p < .001$, $dz = 1.32$,[2] 95% CI [1.04; 1.58] ($M_{low} = -0.97$, $SD_{low} = 1.35$, $M_{med} = -0.01$, $SD_{med} = 1.62$, $M_{high} = 1.59$, $SD_{high} = 1.27$).

Halo effect (mixed-model). For this (and the following) analysis, we estimated a model having attractiveness (C1 and C2) as fixed effects and its random slopes for participant and trait random factors. As expected, the average rating (with negative traits reversed, except vanity-related traits) was higher for high attractive faces compared to low attractive faces, $t(74.71) = 2.74$, $p = .008$ (linear contrast C1; $M_{low} = 2.88$, $SD_{low} = 1.39$, $M_{med} = 2.63$, $SD_{med} = 1.44$, $M_{high} = 3.14$, $SD_{high} = 1.26$). This effect was moderated by target gender, $t(98.27) = 3.31$, $p = .001$, so that a halo effect emerged for female but not for male targets.[3]

Trait selectivity (mixed-models). We added the trait halo relevance to the mixed-model as a fixed effect, as well as its random slope for participants. The interaction between the attractiveness linear contrast (low vs. high attractive faces) and the trait relevance linear contrast (low vs. high dimension relevance) was significant, $t(116.71) = 3.53$, $p < .001$. In other words, the difference in trait ratings for high compared to low attractive faces (i.e., the halo effect) was larger for high-relevance than low-relevance traits. Importantly, when adding trait valence in the model, the observed relevance-based trait selectivity effect was not moderated by trait valence, $t(35.93) = 1.55$, $p = .13$, and remained significant, $t(98.36) = 3.17$, $p = .002$. Moreover, the interaction between attractiveness linear contrast and trait valence was not significant, $t(36.04) = 0.40$, $p = .69$. Results for each trait (i.e., as a function of the trait dimension) and residual contrasts are presented in the Supplementary Materials.

Discussion

Study 1 replicated the classic attractiveness halo effect and its trait selectivity. The attractiveness of a face influenced how it was rated on a series of personality traits and outcomes. This halo effect was larger for traits known to be highly sensitive to attractiveness (i.e., vanity and social competence) than traits known to be less sensitive to that same feature (i.e., integrity and concerns for others; Eagly et al., 1991). Trait selectivity was not explained by differences between traits in terms of valence. Thus, the halo effect – a feature transformation effect implying different source (attractiveness) and target features (e.g., social competence) but identical source and target faces – did not reflect a mere valence-based or "what is beautiful is good" effect.

2 We computed the effect sizes (Cohen's d) for the classic OLS regression analyses ('by-participants' analyses), however, we did not compute the effect sizes for the mixed-model analyses given that there is no clear consensus on this matter (Judd et al., 2017).

3 Simple effects analysis revealed that the halo effect was significant for female targets, $t(123.21) = 44.88$, $p < .001$, with higher ratings for the high as compared to the low attractive faces ($M_{low} = 2.80$, $SD_{low} = 1.39$, $M_{med} = 2.83$, $SD_{med} = 1.37$, $M_{high} = 3.22$, $SD_{high} = 1.21$). However, the halo effect did not reach significance for male targets, $t(113.34) = 0.33$, $p = .74$, that is, high and low attractive faces did not significantly differ ($M_{low} = 2.97$, $SD_{low} = 1.40$, $M_{med} = 2.44$, $SD_{med} = 1.47$, $M_{high} = 3.06$, $SD_{high} = 1.30$). This effect was not observed in Studies 2-6 and therefore is not considered in further detail.

Study 2

In Study 2, we moved from the classic halo effect to a pairing-based halo effect. Rather than presenting a single stimulus, we used a conditioning procedure that paired two stimuli together: a source face (that was either low, medium, or high in attractiveness) and a target face (that was of medium attractiveness). Afterwards, we asked participants to rate the target faces on the same 42 personality traits and outcomes (target features) as in Study 1. Our aim here was to see if the attractiveness of the source face would influence how people evaluate the target face on various trait dimensions.

Following Study 1, we tested whether those effects would be larger for traits high compared to low in halo relevance. If so, this would suggest that the effect is relevance- rather than valence-based. To do so, we used continuous scores of trait halo relevance and trait valence. Trait halo values were computed by taking the linear contrast score (i.e., the difference in ratings for high and low attractive people) for each trait. Of note, the trait halo values were similar to the low relevance/high relevance categories used in Study 1. For instance, the two traits that were the most halo relevant were "popular" and "materialistic" whereas the two less relevant traits were "egoistic" and "intelligent". We then tested whether the pairing-based halo effects increased as a function of this continuous score, that is, whether the pairing-based effect was "relevance-based". We also relied on trait valence (see pilot study in Study 1) to test whether the relevance-based trait selectivity was maintained when trait valence was added to the model. If not, and trait valence moderates the pairing-based halo effect, this would reflect a "valence-based" trait selectivity. Finally, we wanted to know if the halo effect we obtained was sensitive to known moderators of conditioning, that is, people's memory of the stimuli that were paired together (Förderer & Unkelbach, 2014; Stahl & Unkelbach, 2009). We, therefore, assessed at the end of the study whether participants could remember which stimuli were paired during the pairing phase and tested whether memory accuracy was correlated with the magnitude of the pairing-based halo effect.

Method

Participants and design. Our power analysis relied on the halo effect size obtained in Study 1 (dz = 0.39; estimated via OLS regression) to achieve sufficient power (80%). We recruited 200 participants (M_{age} = 23.71, SD_{age} = 3.84, 101 women, 97 men, and 2 participants responding "other"), giving us an 80% likelihood of detecting a minimum effect of dz = 0.18. Participants took part in exchange for £1.50 via Prolific Academic. Inclusion criteria were the same as in Study 1. Following our pre-registered exclusion criteria, we removed one participant who had zero variance in their personality traits/outcomes ratings. A 3 (*Attractiveness*: low vs. medium vs. high) x Continuous (*Trait Halo Relevance*: from -0.74 to +1.43) x Continuous (*Trait Valence*: from -2.10 to +2.61) x 2 (*Pairing Memory*: correct vs. incorrect) x 2 (*Target Gender*: male vs. female) design was employed with the final two factors being between-participants.

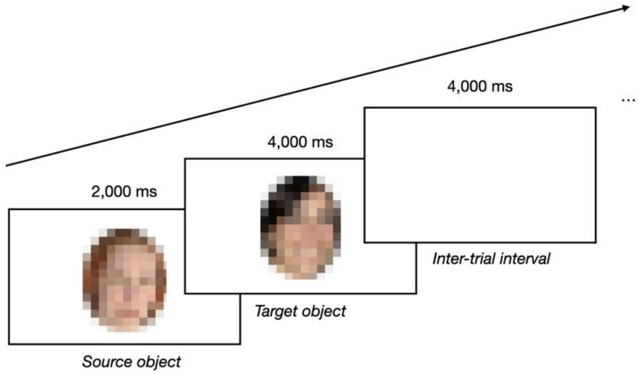

Figure 1. Time course of a trial in the pairing phase

Note. A source face (low vs. medium vs. high on attractiveness) was displayed on the screen for 2000 ms, followed by a target face (medium on attractiveness) for 4000 ms. The original face pictures used in our procedure have been replaced by pixelated face examples (i.e., not used in any of our studies) because the license agreement of the 10k US Adult Faces Database does not allow for reproduction or modification of the face pictures used in the study. The inter-trial interval duration was 4000 ms. Participants' task was to remain focused and to watch the faces appearing on the screen.

Materials. The same 12 faces were used as source faces as in Study 1. As target faces, six new faces were selected (3 male and 3 female) from the same face database (Bainbridge et al., 2013). Selected faces were white, unknown (i.e., not celebrities), most of them belonged to the 20-30 years old category, and scored relatively average on physical attractiveness (M = 5.95, SD = 0.70).

We relied on the same 42 personality traits and outcomes as in Study 1. We used the scores of trait halo relevance obtained in Study 1 to test for the relevance-based trait-selectivity (i.e., the difference in trait ratings for high compared to low attractive faces) and trait valence scores obtained in the pilot study to test for the alternative valence-based trait-selectivity.

Procedure. After providing their consent, participants received the same general instruction as in Study 1. However, this time, they were also informed that before judging the faces, they would see pairs of faces. They thus had to remain focused and watch these faces carefully.

Pairing phase. During each trial, two faces were presented sequentially one after the other. One of these faces was either low, medium, or high in attractiveness (source face), and the other face was of medium attractiveness (target face). For each participant, stimuli were selected from a pool of 12 source and 6 target faces with two selection conditions: (a) the three selected source faces had to cover the three levels of attractiveness (low vs. medium vs. high), and (b) the three selected target faces had to match in gender with the source faces (i.e., all men or all women). Each participant was thus presented with three pairs of faces. Each pair was presented five times, resulting in 15 trials per participant. Within each pair, the source face was presented for 2000 ms, immediately followed by the target face for 4000 ms (see Figure 1). The inter-trial interval between two pairs was 4000 ms.

Rating phase. Participants evaluated the three target faces on the same 42 personality traits and outcomes (and both source and target faces in terms of their attractiveness) as in Study 1. Again, faces were displayed at the top of the screen, with the rating scales below.

Pairing memory, demographic questions, and exploratory questions. During the memory task, each target face was presented one at a time at the top, along with the three source faces at the bottom of the screen. Participants were asked to indicate which source face had been paired with a given target face during the pairing phase.

Thereafter participants provided demographic information (age, gender, English fluency), optional comments about the study, and answered a series of exploratory questions related to perceived demand awareness, compliance, source feature awareness, and influence. Because of their exploratory nature, these questions will not be discussed further (for more information, see Supplementary Materials). Participants were then thanked and debriefed.

Results

Participants who made at least one error when reporting their pairing memory were coded as having an incorrect memory. We used the same orthogonal contrast codes for source face attractiveness as in Study 1. Trait halo relevance and valence were centred on zero, and pairing memory was contrast coded (incorrect: -0.5; correct: +0.5). We again reversed the ratings for negatively valenced traits except for vanity-related traits. Mixed-models were the same as the ones in Study 1, except that we added memory as a fixed effect when testing for its moderation.

Attractiveness ratings (OLS regression). Participants judged high attractive source faces as more attractive than low attractive source faces, $t(197) = 16.70$, $p < .001$, $dz = 1.19$, 95% CI [1.01; 1.37] ($M_{low} = -1.22$, $SD_{low} = 1.48$, $M_{med} = -0.50$, $SD_{med} = 1.74$, $M_{high} = 1.25$, $SD_{high} = 1.52$). However, the perceived attractiveness of the targets paired with high attractive faces did not differ from the targets paired with low attractive faces, $t(197) = 0.12$, $p = .90$, $dz < 0.01$, 95% CI [-0.13; 0.15] ($M_{low} = 0.34$, $SD_{low} = 1.45$, $M_{med} = 0.07$, $SD_{med} = 1.50$, $M_{high} = 0.32$, $SD_{high} = 1.52$).[4]

Pairing-based halo effect (mixed-model). The pairing-based halo effect (i.e., the linear effect of the source faces' attractiveness on the target faces' ratings) was not significant, $t(154.91) = 1.30$, $p = .19$ ($M_{low} = 3.08$, $SD_{low} = 1.17$, $M_{med} = 3.07$, $SD_{med} = 1.22$, $M_{high} = 3.19$, $SD_{high} = 1.19$). This effect was moderated by gender, $t(195.80) = 2.88$, $p = .004$, such that the pairing-based halo effect emerged only for female (not for male) participants.[5]

Trait selectivity: relevance- or valence-based? (mixed-model). On the one hand, the interaction between attractiveness linear contrast (low vs. high attractive source faces) and trait halo relevance was not significant, $t(196.74) = 1.79$, $p = .07$. Contrary to our predictions, the pairing-based halo effect descriptively decreased (rather than increased) as a function of increases in trait halo relevance (see Figure 2, *left panel*). This observed trait selectivity effect was not moderated by trait valence, $t(37.99) = 0.96$, $p = .34$, and remained non-significant when trait valence was added to the model, $t(194.60) = 1.13$, $p = .26$.

On the other hand, the pairing-based halo effect was moderated by trait valence, $t(37.99) = 3.16$, $p = .003$, such that the effect was significant for high trait valence (-1*SD*), $t(146.80) = 2.89$, $p = .004$, but not for low trait valence (+1*SD*), $t(136.00) = 0.09$, $p = .93$ (see Figure 2, *right panel*). Taken together, these findings indicate that the observed trait selectivity was more valence- than relevance-based.[6]

Moderation by pairing memory (mixed-model). Pairing memory did not moderate the pairing-based halo effect ($N_{\text{correct}} = 82$, $N_{\text{incorrect}} = 116$), $t(196.20) = 0.99$, $p = .32$, nor did it influence the relevance-based , $t(207.40) = 1.41$, $p = .16$, or the valence-based trait selectivity, $t(23670.00) = 1.17$, $p = .24$. The overall interaction between the attractiveness linear contrast, trait halo relevance, trait valence, and the pairing memory condition was not significant, $t(23670.00) = 0.77$, $p = .44$.

Discussion

An overall pairing-based halo effect did not emerge in Study 2. This effect was not moderated by trait halo relevance and therefore did not seem to be relevance-based. Instead, the pairing-based effect seemed to be valence-based, that is, moderated by trait valence: the more attractive the source face, the more positively evaluated is the target face (i.e., higher ratings on highly positive traits).

One possible explanation for the observed valence-based trait selectivity is that the source faces were processed in terms of their valence rather than attractiveness. Attractive faces could have been encoded as positive (or liked) rather than attractive (and low attractive faces as negative or disliked). This may have influenced subsequent judgements

4 Correlations between source and target attractiveness ratings for each study can be found in the Supplementary Materials.

5 The halo effect was significant for female participants, $t(198.95) = 2.95$, $p = .004$ ($M_{low} = 2.96$, $SD_{low} = 1.21$, $M_{med} = 2.98$, $SD_{med} = 1.26$, $M_{high} = 3.12$, $SD_{high} = 1.26$), but did not reach significance for male participants, $t(200.45) = 0.97$, $p = .33$ ($M_{low} = 2.98$, $SD_{low} = 1.23$, $M_{med} = 2.89$, $SD_{med} = 1.26$, $M_{high} = 2.92$, $SD_{high} = 1.26$).

6 Given its similarity with outcomes generally assessed in EC research, we tested if the pairing-based effect emerged for the 'likability' item. This effect did not emerge, $t(197) = 1.48$, $p = .14$, $dz = 0.11$, 95% CI [-0.03; 0.25] ($M_{low} = 3.19$, $SD_{low} = 1.21$, $M_{med} = 3.20$, $SD_{med} = 1.21$, $M_{high} = 3.37$, $SD_{high} = 1.27$). The absence of this effect could be due to methodological reasons (i.e., the 'likability' item does not directly assess whether *participants* like the target but whether it is likable in general) and statistical reasons (i.e., because using a single item decreases power as compared to testing variation on a continuous score).

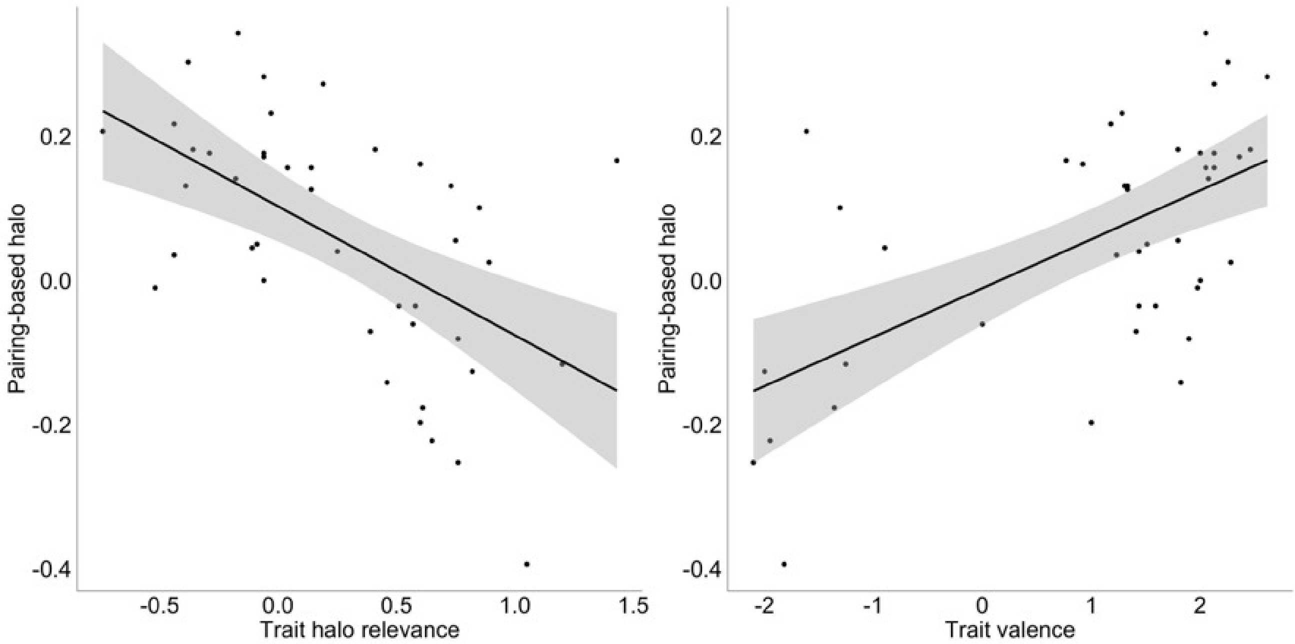

Figure 2. Pairing-based halo effect (High-Low scores) at the trait level, as a function of trait halo relevance (left panel) and trait valence (right panel)

Note. Grey areas represent the 95% confidence intervals. The pairing-based halo effect was not significantly related to trait halo relevance (*left panel*). It was, however, positively related to trait valence (*right panel*) so that the higher the trait valence, the larger the pairing-based effect.

about the target face so that the more attractive the source face, the more positive the ratings on the target face.

This unexpected result echoes what is sometimes observed in conditioning studies. When the source face possesses multiple attributes – as is the case with faces, where valence could be considered as only one of many attributes – the attribute that is conditioned depends on participants' focus (e.g., Gast & Rothermund, 2011). One way to increase the impact of attractiveness is to make this feature salient. In conditioning studies, making salient the target feature (e.g., athleticism) strongly contributes to producing a pairing effect on that specific attribute (Förderer & Unkelbach, 2014). We, therefore, set out to make the source feature of attractiveness salient via instructions in the following study.

Study 3

Study 3 was similar to the previous study but with several notable changes. First, we made the source feature of attractiveness salient via instructions. Participants were now told explicitly how attractive the source faces were according to other people. Second, we asked participants to remember which faces were paired together during the pairing phase (i.e., we oriented participants' attention to the stimuli being paired), which is known to increase conditioning effects and could thus also increase pairing-based halo effects (e.g., Förderer & Unkelbach, 2014; Gast & Rothermund, 2011). Third, we used a simultaneous (instead of sequential) pairing procedure, which also leads to larger conditioning effects and might therefore also increase pairing-based halo effects (e.g., Stahl & Heycke, 2016).

Method

Participants and design. The same sample size estimation, inclusion criteria, and design were used as in Study 2. The only difference was that the target face position (left vs. right) was counterbalanced across participants. We recruited 201 participants (M_{age} = 24.06, SD_{age} = 3.83, 112 women, 87 men, 2 participants responding "other") who took part in exchange for £1.75.

Materials and Procedure. Similar materials and procedures were used as in Study 2 with a few exceptions. First, instructions were added to make attractiveness a salient feature of the source faces. Participants were told that they would see three pairs of faces and the people they would have to judge (target faces) later in the study would be presented on the left (vs. right). Each source face was then presented together with information about its perceived attractiveness levels (i.e., that it was "generally described by others [friends, family, and strangers] as being relatively low [vs. average vs. high] on physical attractiveness"). To ensure that attractiveness remained salient, source faces were presented together with attractiveness labels ("low attractiveness" vs. "average attractiveness" vs. "high attractiveness") during the pairing phase.

Second, immediately before the pairing phase, participants were told to "look carefully at all the faces (i.e., both faces on the left and faces on the right)" and to "try to remember which faces were presented together in a pair". This was meant to increase the chances of accurate pairing memory. Finally, the pairing phase used a simultaneous presentation of faces. Faces were presented together on the screen for 2,500 ms, and each pair were separated by an inter-trial interval of 1,000 ms (see Figure 3).

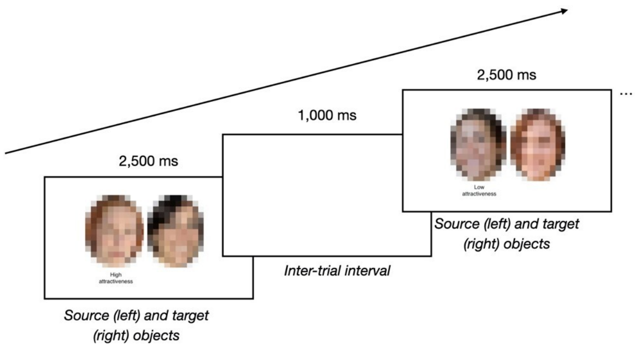

Figure 3. Time Course of two trials during the pairing phase

Note. The source face (low vs. medium vs. high on attractiveness) and target face (medium on attractiveness) were simultaneously presented onscreen for 2500 ms, followed by an inter-trial interval of 1000 ms. The original face pictures used in our procedure have been replaced by pixelated face examples (i.e., not used in any of our studies) because the license agreement of the 10k US Adult Faces Database does not allow for reproduction or modification of the face pictures used in the study. Source faces were presented together with an attractiveness label under them.

Results

The same coding and transformations were used as in Study 2. Two participants who had zero variance in their personality traits/outcomes ratings were excluded. Target face position did not moderate main effects in any study and was therefore removed from subsequent analyses. Mixed-models were the exact same as in Study 2.

Attractiveness ratings (OLS regression). Participants judged high attractive source faces as more attractive than low attractive source faces, $t(198) = 18.00$, $p < .001$, $dz = 1.28$, 95% CI [1.09; 1.47] ($M_{low} = -1.34$, $SD_{low} = 1.39$, $M_{med} = -0.37$, $SD_{med} = 1.57$, $M_{high} = 1.14$, $SD_{high} = 1.46$). The target faces paired with high attractive faces were not rated as more attractive than the target faces paired with low attractive faces, $t(198) = 1.26$, $p = .21$, $dz = 0.09$, 95% CI [-0.05; 0.23] ($M_{low} = 0.17$, $SD_{low} = 1.47$, $M_{med} = 0.18$, $SD_{med} = 1.38$, $M_{high} = 0.33$, $SD_{high} = 1.39$).

Pairing-based halo effect (mixed-model). An overall pairing-based halo effect did not emerge, $t(91.82) = 1.31$, $p = .19$ (linear contrast C1; $M_{low} = 2.92$, $SD_{low} = 1.27$, $M_{med} = 2.91$, $SD_{med} = 1.23$, $M_{high} = 2.99$, $SD_{high} = 1.23$).

Trait selectivity: relevance- or valence-based? (mixed-model). On the one hand, the interaction between attractiveness linear contrast and trait halo relevance was significant and in the predicted direction, $t(205.64) = 3.98$, $p < .001$. For traits high in halo relevance (+1*SD*) the pairing-based halo effect was significant and in the expected direction, $t(203.29) = 4.00$, $p < .001$. For traits low in halo relevance (-1*SD*) the pairing-based halo effect was also significant but in the opposite direction, $t(193.63) = 2.57$, $p = .01$ (see Figure 4, *left panel*).

When adding trait valence in the mixed-model, this relevance-based trait selectivity effect was not significantly moderated by trait valence, $t(38.20) = 0.41$, $p = .68$, and remained significant, $t(207.60) = 3.62$, $p < .001$. The pairing-based halo effect was not moderated by trait valence, $t(38.20) = 1.76$, $p = .09$ (see Figure 4, *right panel*).

Moderation by pairing memory. The pairing-based halo effect was significantly moderated by pairing memory ($N_{correct} = 131$, $N_{incorrect} = 68$), $t(196.59) = 3.14$, $p = .002$. Simple effect analyses revealed that participants who had a correct memory of the pairings produced a significant pairing-based halo effect, $t(192.55) = 3.27$, $p = .001$ ($M_{low} = 2.88$, $SD_{low} = 1.28$, $M_{med} = 2.90$, $SD_{med} = 1.21$, $M_{high} = 3.04$, $SD_{high} = 1.22$), whereas no such effect emerged for those classified as having incorrect memory, $t(204.96) = 1.37$, $p = .17$ ($M_{low} = 3.00$, $SD_{low} = 1.25$, $M_{med} = 2.94$, $SD_{med} = 1.27$, $M_{high} = 2.91$, $SD_{high} = 1.25$). Pairing memory also moderated the relevance-based trait selectivity, $t(204.30) = 3.21$, $p = .002$, with a larger trait selectivity for participants having a correct (vs. incorrect) memory of the pairing. This moderation was not significant for the valenced-based trait selectivity, $t(23790.00) = 1.00$, $p = .32$. The overall interaction between the attractiveness linear contrast, trait halo relevance, trait valence, and the pairing memory condition was not significant, $t(23790.00) = 0.53$, $p = .59$.

Discussion

Although the pairing-based halo effect in Study 3 was not significant – on average – it did emerge as expected for target traits known to be highly sensitive to the source feature of attractiveness. Of note, the fact that the average pairing-based halo effect did not emerge is not surprising as we expected the direction of this effect to vary as a function of the target feature (i.e., to be trait selective), based on the results observed in Study 1 (e.g., to be in a positive direction for some sociability-related traits and in a negative direction for some integrity-related traits). Importantly, the relevance-based trait selectivity remained significant when controlling for the trait valence. This pattern of results contrasts sharply with what we observed in Study 2. A key difference between Studies 2 and 3 which may explain the contrasting results is the salience of attractiveness. When this source feature was made salient (as in Study 3), the trait selectivity was relevance-based. When not salient (as in Study 2), the trait selectivity was valence-based. However, the two studies also differed in other ways (e.g., sequential vs. simultaneous presentation of pairs). To examine if our findings were primarily driven by attractiveness salience, we directly manipulated this factor in Study 4.

Study 4

Study 4 had two main goals. First, we set out to replicate our previous findings (i.e., the moderating role of trait halo relevance and pairing memory when attractiveness is salient). Second, we aimed to directly compare pairing-based halo effects when source attractiveness was non-salient versus salient: we manipulated the salience of attractiveness between participants with the same pairing procedure as before. We expected a larger pairing-based halo effect in the salient compared to the non-salient condition. We also expected the nature of the trait selectivity to differ: when attractiveness is made salient (i.e., when the source faces are processed on attractiveness), the effect

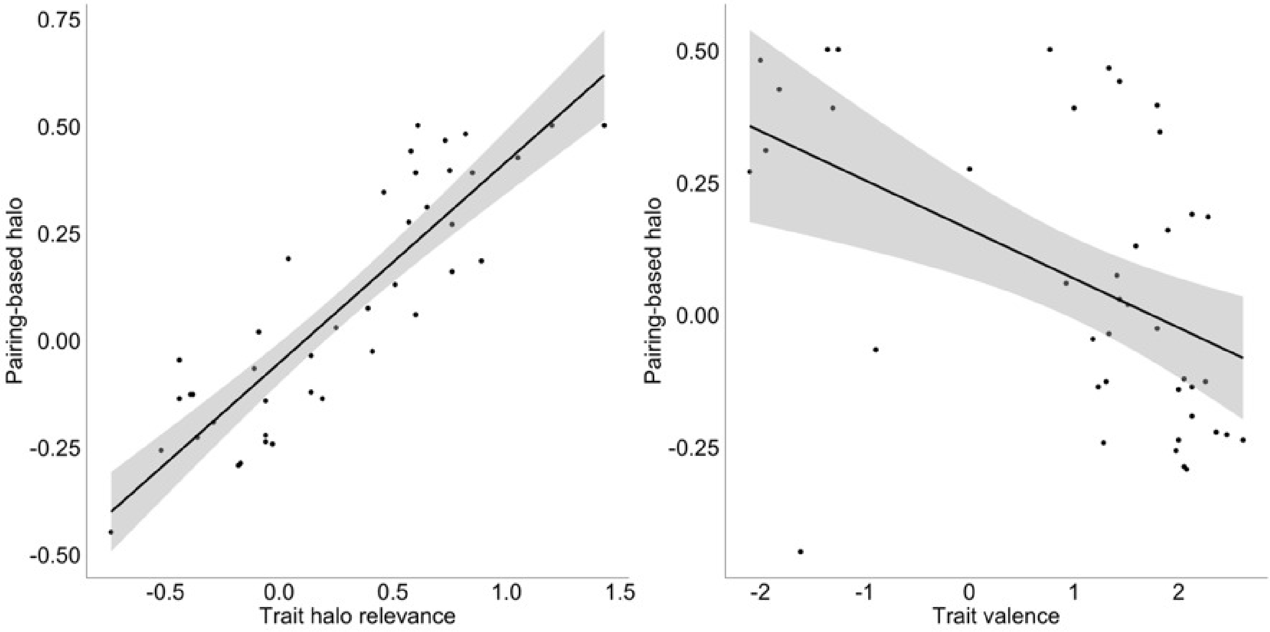

Figure 4. Pairing-based halo effect (High-Low scores) at the trait level, as a function of trait halo relevance (left panel) and trait valence (right panel)

Note. Grey areas represent the 95% confidence intervals. The pairing-based halo effect was positively related to trait halo relevance (*left panel*) so that the higher the trait relevance, the larger the pairing-based effect. It was not significantly related to trait valence (*right panel*).

should be relevance-based (replicating results of Study 3), whereas when attractiveness is not salient (i.e., when the source faces are processed on valence), the trait selectivity should be valence-based (replicating results of Study 2).

Method

Participants and design. To estimate our sample size, we relied on Studies 2 and 3. Specifically, we merged the studies into one dataset and computed the effect size of the pairing-based halo by salience effect (d = 0.42 across the studies). A sample of 298 participants provided 95% power. Nevertheless, to allow for exclusions and procedural changes, we recruited 401 participants (M_{age} = 24.36, SD_{age} = 4.07, 227 women, 170 men, and 4 participants responding "other") who took part in exchange for £.1.75. Inclusion criteria were the same as before. The design of Study 4 was similar to Study 3's except for the addition of a between-subject manipulation of the salience of attractiveness.

Materials and Procedure. The materials and procedure in the "salience" condition were the same as in Study 3. Participants received information about the source faces' attractiveness (i.e., how attractive these people are generally perceived to be), along with the corresponding label (low vs. average vs. high attractiveness). Instructions in the "no salience" condition were almost identical but never implied any reference to physical attractiveness. During the instruction phase of the no salience condition, the source faces were displayed, and participants were instructed to look carefully at the faces. The pairing phase and other measures were the same as in Study 3.

Results

We used the same coding and transformation as before. We excluded 7 participants with zero variance in their personality traits/outcomes ratings. The salience variable was contrast coded (no salience: -0.5; salience: +0.5). Results are reported with the same structure as before, adding those related to the attractiveness salience moderation as the last section. Mixed-models were the same as previously, except that we added salience as a fixed effect when testing for its moderation.

Attractiveness ratings (OLS regression). Participants judged high attractive source faces as more attractive than low attractive source faces, $t(393)$ = 26.22, p < .001, dz = 1.32, 95% CI [1.19; 1.46] (M_{low} = -1.21, SD_{low} = 1.57, M_{med} = -0.23, SD_{med} = 1.63, M_{high} = 1.45, SD_{high} = 1.27). Following the pairing procedure, target faces paired with attractive source faces were rated as more attractive than the target faces paired with low attractive source faces, $t(393)$ = 2.57, p = .01, dz = 0.13, 95% CI [0.03; 0.23] (M_{low} = 0.21, SD_{low} = 1.58, M_{med} = 0.38, SD_{med} = 1.55, M_{high} = 0.47, SD_{high} = 1.51). This effect significantly varied as a function of the salience of the source feature, $t(392)$ = 2.13, p = .03, dz = 0.22, 95% CI [0.02; 0.41], so that it was larger in the salience condition (M_{low} = 0.08, SD_{low} = 1.58, M_{med} = 0.39, SD_{med} = 1.57, M_{high} = 0.56, SD_{high} = 1.68) than in the no salience condition (M_{low} = 0.34, SD_{low} = 1.57, M_{med} = 0.37, SD_{med} = 1.52, M_{high} = 0.38, SD_{high} = 1.31).

Pairing-based halo effect (mixed-model). An overall pairing-based halo effect emerged, $t(190.50)$ = 2.39, p = .02 (linear contrast C1; M_{low} = 2.92, SD_{low} = 1.27, M_{med} = 2.92, SD_{med} = 1.23, M_{high} = 2.99, SD_{high} = 1.23), meaning that target faces paired with high attractive faces received a higher average score than target faces paired with low attractive faces.

Trait selectivity: relevance- or valence-based? (mixed-model). On the one hand, the interaction between the attractiveness linear contrast and trait halo relevance was significant and in the expected direction, $t(377.88) = 2.44$, $p = .01$. The pairing-based halo effect increased as a function of trait halo relevance, and was significant for traits high in halo relevance (+1*SD*), $t(361.97) = 3.48$, $p < .001$, but not for traits low in halo relevance (-1*SD*), $t(357.65) = 0.53$, $p = .59$. On the other hand, the relevance-based trait selectivity was not moderated by trait valence, $t(37.98) = 0.18$, $p = .86$, and it remained significant even after trait valence was added to the model, $t(377.49) = 2.35$, $p = .02$. Finally, the interaction between attractiveness linear contrast and trait valence was not significant, $t(37.98) = 0.26$, $p = .80$.

Moderation by pairing memory (mixed-model). The pairing-based halo effect was moderated by pairing memory ($N_{correct} = 254$, $N_{incorrect} = 140$), $t(392.18) = 2.13$, $p = .03$. Participants who had a correct memory of the pairings produced a significant pairing-based halo effect, $t(392.14) = 3.48$, $p < .001$ ($M_{low} = 2.91$, $SD_{low} = 1.28$, $M_{med} = 2.97$, $SD_{med} = 1.23$, $M_{high} = 3.03$, $SD_{high} = 1.22$), whereas those who had an incorrect memory failed to produce an effect, $t(382.36) = 0.06$, $p = .95$ ($M_{low} = 2.94$, $SD_{low} = 1.26$, $M_{med} = 2.98$, $SD_{med} = 1.26$, $M_{high} = 2.94$, $SD_{high} = 1.23$). The relevance-based trait selectivity was also moderated by the pairing memory, $t(406.60) = 2.64$, $p = .009$ (with a larger trait selectivity for participants having a correct rather than incorrect memory), whereas this was not the case for valence-based trait selectivity, $t(47150.00) = 1.91$, $p = .06$. The interaction between the attractiveness linear contrast, trait halo relevance, trait valence, and pairing memory was not significant, $t(47150.00) = 0.15$, $p = .88$.

Moderation by attractiveness salience (mixed-model). The pairing-based halo effect was not moderated by the salience manipulation, $t(392.40) = 0.70$, $p = .48$. Our analyses did reveal, however, a four-way interaction between the attractiveness linear contrast, trait halo relevance, trait valence, and salience, $t(47150.00) = 2.06$, $p = .04$, suggesting that the type of observed trait selectivity of the halo effect varied as a function of the salience condition (see Figure 5).

The three-way interaction between the attractiveness linear contrast, trait halo relevance, and salience was significant, $t(407.00) = 2.85$, $p = .005$. Specifically, the interaction between attractiveness and trait halo relevance was larger and significant in the salience condition, $t(403.00) = 3.70$, $p < .001$, as compared to the no salience condition, for which the interaction was not significant, $t(403.20) = 0.35$, $p = .73$. Moreover, the three-way interaction between the attractiveness linear contrast, trait valence, and salience was also significant, $t(47150.00) = 3.23$, $p = .001$, but in the opposite direction. Specifically, the interaction between attractiveness and trait valence was larger and just significant in the no salience condition, $t(126.90) = 1.98$, $p = .05$ (rounded up from $p = .049$),[7] as compared to the salience condition, for which the effect was also significant, $t(118.00) = 2.32$, $p = .02$, but in the opposite direction (i.e., decrease in the pairing-based halo effect as trait valence increased).

Discussion

Study 4 revealed three moderators of pairing-based halo effects: trait halo relevance, pairing memory, and source feature salience. Taken together, the results of Studies 2-4 suggest that, under certain conditions, the attractiveness halo effect can be established via a pairing procedure. Moderators typical of the attractiveness halo effect (relevance-based trait selectivity) and of conditioning effects (pairing memory, source feature salience) strongly moderate these effects, with source feature salience serving as a particularly interesting moderator. When the attractiveness is not salient, the source face influences how people respond to the target face on the basis of valence rather than attractiveness, whereas the opposite is true when attractiveness is salient.

In Studies 5-6, we set out to extend and generalise our findings to another type of halo effect (i.e., the health halo effect). Our initial goal was to replicate the classic health halo effect and its trait selectivity (Study 5) and then examine if we could also obtain a pairing-based health halo effect (Study 6). We also tested the moderating role of trait halo relevance (vs. trait valence) and pairing memory in the pairing-based design.

Studies 5-6: Health Halo Effects

Study 5

We attempted to replicate the health halo effect by manipulating food labels and the information they communicate about the fat content of a product (low vs. medium vs. high fat; Ebneter et al., 2013; Westcombe & Wardle, 1997). Products with lower fat content (source feature) should be rated more positively on a series of target features (e.g., healthier) than products higher in their fat content. We also examined for evidence of a relevance-based trait selectivity (Lee et al., 2013; Orquin & Scholderer, 2015). Similar to Studies 1-4, we relied on Study 5 to create continuous scores of trait halo relevance (i.e., to what extent a trait produces a health halo effect) and trait valence (i.e., to what extent a trait is perceived as being positive or negative) for later use in Study 6.

7 Technically, the observed effect was significant because the *p*-value was below the conventional threshold of significance, $\alpha = 0.05$, that we established *a priori* (i.e., before collecting and analysing data). Pending replication, however, this effect should be interpreted with caution given that it is fairly close to the critical threshold.

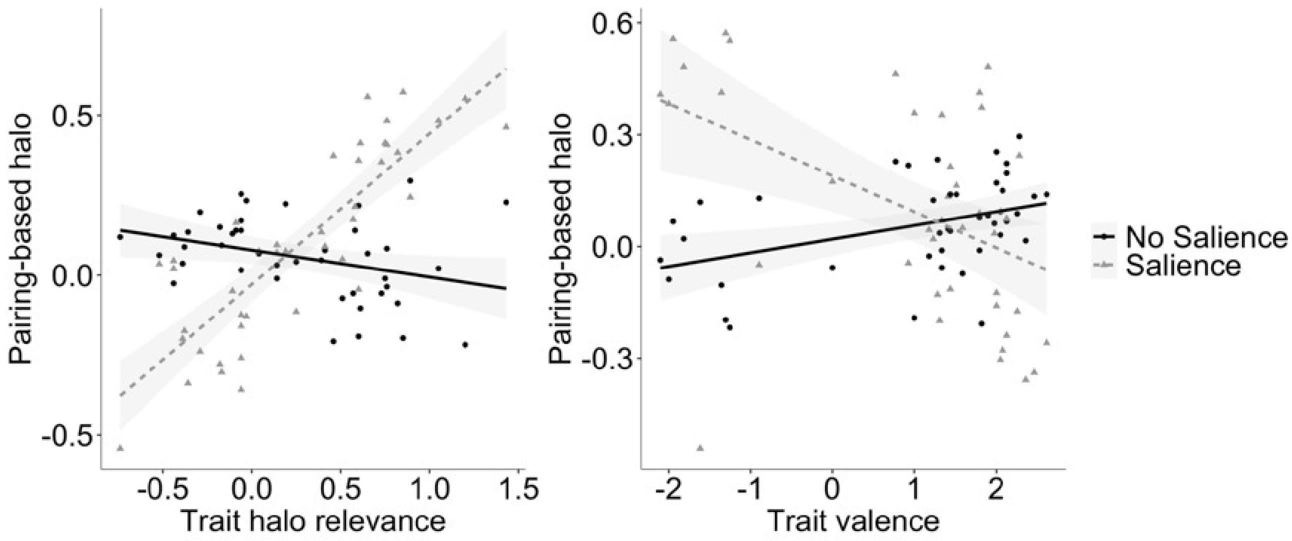

Figure 5. Pairing-based halo effect (High-Low scores) at the trait level, as a function of the salience, trait halo relevance (left panel), and trait valence (right panel)

Note. Grey areas represent the 95% confidence intervals. As can be seen on the left panel, the pairing-based effect was positively related to trait halo relevance in the salience condition, whereas this relation was not significant in the no salience condition. On the right panel, the pairing-based effect was negatively related to trait valence in the salience condition, whereas this relation was positive in the no salience condition.

Method

Participants and design. Previous studies examining the health halo effect reported effect sizes ranging from small to large (d = 0.18-0.70; Andrews et al., 2009; Ebneter et al., 2013). Based on available resources, we recruited a sample of 106 participants (M_{age} = 34.38, SD_{age} = 15.97, 50 women, 55 men, and 1 participant who did not select a gender), giving us 80% of power to detect a minimum effect size of dz = 0.25. Participants spoke Italian as a first language and took part voluntarily in the study. A 3 (*Fat Content*: low vs. medium vs. high) x 3 (*Trait Relevance*: low vs. medium vs. high) x continuous (*Trait Valence*: from 1 to 7) within-participants design was employed.

Materials. Coloured photographs of three packs of cookies served as stimuli. Information that could lead participants to identify the brand of cookies was removed, and the generic name "cookies" (originally "Biscotti" in Italian) was added to the cookie pack image. Twenty-one traits used in the health halo literature were selected as target features. Based on previous results of Oostenbach et al. (2019), we classified these traits into three categories of trait relevance regarding the source feature of fat content: 1) fibre content, healthy, and light were "highly relevant" (i.e., they were expected to produce the largest halo effect), 2) tasty, digestible, and nourishing were "medium relevant" (i.e., they were expected to produce an intermediate halo effect), and 3) fresh, environmentally friendly, organic, expensive, safe to eat, of quality, satiating, trustworthy, attractive, appetizing, spicy, fragrant, sweet, and salty were "low relevant" (i.e., they were expected to produce the lowest health halo effect or no effect at all).[8]

Procedure. The experiment was programmed in Qualtrics and administered online. After providing their consent, participants were told that the study was concerned with food perception and that our aim was to measure how accurate they were in judging food qualities. They were also informed that their performance would be compared with individuals who received special training (nutritionist students) but that people without training can often be as good as their trained counterparts.

Participants were then randomly presented with the three products, one at a time, and were asked to evaluate each product along the 21 traits (all scales ranged from 1 = *not at all* to 7 = *very much*). Each product was displayed at the top of the screen, with the rating scales below. Once all evaluations for one product were complete, they moved on to the next one. Products were presented with a description indicating their level of fat content: "These cookies have a low (vs. medium vs. high) fat content (2.3g [vs. 12.3g vs. 22.3g] /100g, 3.07% [vs. 16.4% vs. 29.7%])". The label attached to each package was counterbalanced between participants. The order of the 21 traits was randomly established prior to the study and held constant across participants, with the final two traits always being calorie content and fat content.

Then, participants reported their perceived trait valence by rating each trait in terms of its valence on a 7-point scale (from -3 [*extremely negative*] to +3 [*extremely positive*]). The traits were presented in the same order as during the

8 Alternatively, appetizing and fragrant could have been classified into the medium relevance category if one considers that they are linked to tastiness. This alternative classification does not influence the significance of the presented results.

rating phase. Participants then answered a series of exploratory questions related to perceived demand awareness, demand compliance, source feature awareness, and influence (see Supplementary Materials). Finally, they provided demographic information (age, gender), indicated whether their data could be trusted (self-reported single item indicator, SRSI; Meade & Craig, 2012), after which they were thanked and debriefed.

Results

Participants who indicated that we should not use their data were removed (N = 6). We used orthogonal contrast codes for the fat content and trait relevance variables (linear contrast C1: low = -1/2, medium = 0, high = 1/2; quadratic contrast C2: low = -1/3, medium = 2/3, high = -1/3). Again, we focused on the linear contrast code opposing low and high (i.e., the low vs. high-fat content; the low vs. high relevant traits). The results for the quadratic contrasts are presented in Supplementary Materials. We reversed the ratings for the negatively valenced trait 'expensive' so that for all ratings, a higher score implied a more positive judgement. Below, we report the results for the fat content ratings, the halo effect (i.e., the average rating score for the 21 traits), and its trait-selectivity. Mixed-models were the same as the ones used in Study 1.

Fat content ratings (OLS regression). We computed an average score of calorie and fat content ratings. In line with our manipulation, participants judged high fat products as having more calories and fat than the low fat products, $t(99) = 11.23$, $p < .001$, $dz = 1.13$, 95% CI [0.87; 1.38] ($M_{low} = 3.54$, $SD_{low} = 1.67$, $M_{med} = 4.64$, $SD_{med} = 1.21$, $M_{high} = 5.76$, $SD_{high} = 1.23$).

Halo effect (mixed-model). On average, target trait ratings were not significantly lower for high fat products compared to low fat products (linear contrast C1), $t(26.57) = 1.92$, $p = .07$ ($M_{low} = 4.12$, $SD_{low} = 1.60$, $M_{med} = 3.92$, $SD_{med} = 1.59$, $M_{high} = 3.85$, $SD_{high} = 1.71$).

Trait selectivity (mixed-model). We tested whether the halo effect (i.e., the linear effect of fat content) was moderated by trait relevance. The interaction between the fat content linear contrast (low vs. high fat content) and trait relevance linear contrast (low vs. high relevance) was significant, $t(21.96) = 3.13$, $p = .005$. The halo effect was larger for the high than for the low relevance traits (see Figure 6). Results for each trait and the residual contrasts are in Supplementary Materials. This interaction was not moderated by trait valence, $t(818.30) = 0.72$, $p = .47$, and remained significant even after this variable was added to the model, $t(22.62) = 2.98$, $p = .007$. Trait valence did not significantly moderate the halo effect, $t(1559) = 1.95$, $p = .05$.

Discussion

We produced a health halo effect that was trait selective and consistent with past work in the halo literature (i.e., that larger halo effects emerge for high than for low relevant traits; Oostenbach et al., 2019). Interestingly, the observed relevance-based trait selectivity remained when

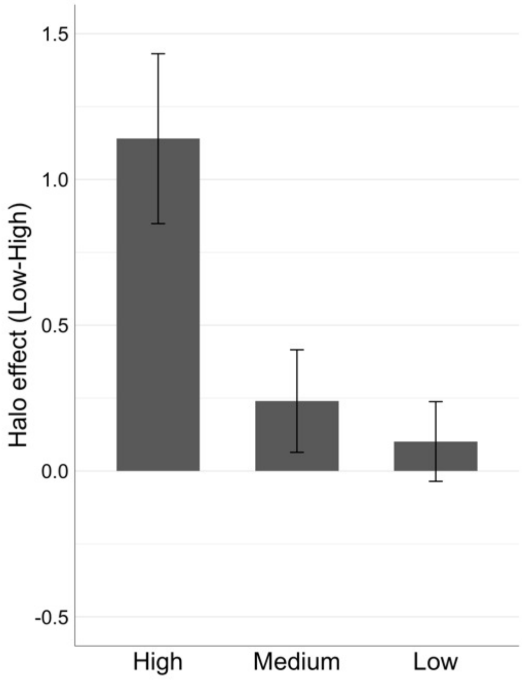

Figure 6. Halo effect (Low-High scores) as a function of the trait relevance classification (high vs. medium vs. low)

Note. Error bars represent the 95% confidence intervals.

controlling for trait valence. In the following study, we set out to produce a pairing-based health halo effect.

Study 6

In our final study, the design was broadly similar to that outlined in Study 3. More specifically, the source feature (fat content) was salient, we used a simultaneous pairing procedure, and we asked participants to attend and memorize the paired stimuli. As before, we relied on continuous measures of trait halo relevance (i.e., for each trait, the difference in ratings for high- and low-fat products in Study 5) and trait valence (i.e., for each trait, the mean valence rating collected in Study 5). This enabled us to test whether the pairing-based health halo effect would be trait selective in the same way as the standard health halo effect (i.e., relevance-based) or whether it would depend on the valence of the traits (i.e., valence-based). We also examined the moderating impact of pairing memory on these effects. More exploratory, we tested if the memory of the source feature (i.e., whether participants remembered correctly vs. incorrectly the fat content of the source product) influenced the pairing-based effect.

Method

Participants and design. We relied on the same power analysis as in Study 4. However, to be more conservative, we recruited 148 participants ($M_{age} = 26.23$, $SD_{age} = 6.82$, 64 women, 80 men, and 2 participants preferred not to say),[9] giving us 80% of power to detect a minimum effect size of *dz* = 0.20. Participants took part via Prolific Academic in exchange for £1.50, spoke Italian as a first language, and had an approval rate of at least 70%. A 3 (*Fat Content*: low vs. medium vs. high) x continuous (*Trait Halo Relevance*: from -1.46 to +0.67) x continuous (*Trait Valence*: from 1 to 7) x 2 (*Pairing Memory*: correct vs. incorrect) x 2 (*Source Feature Memory*: correct vs. incorrect) design was employed with the two final variables being between participants.

Materials. We used the same 3 food pictures as in Study 5 along with 3 new pictures. Assignment of the cookie packs to serve as source products (with low vs. medium vs. high-fat content) or target products (i.e., without any fat content description) was randomised. The same 21 traits were used to assess target judgements as in Study 5. The trait halo relevance was based on the scores of health halo observed for each trait in Study 5 (i.e., the difference of rating between high and low-fat content for each trait).

Procedure. The experiment was programmed in Inquisit. After providing consent, participants received the same general instructions as in Study 5. They were also informed that products would be presented as pairs, that they would see three pairs of images, and in each case, they would see one product that they would later need to judge. They were then shown the source products that would be presented on the right (vs. left) side of the screen, along with the same fat content description as in Study 5.

Pairing phase. The same simultaneous pairing procedure was used as in Study 3. Each pair consisted of a source product that varied in its fat content (low vs. medium vs. high) and a target product that had no mention of its fat content. The position of the target product on the screen (left vs. right) was randomised between participants – note that this parameter was not recorded in the data, so we could not test for its potential effect. The pairing phase consisted of three pairs of stimuli that covered the three levels of fat content. Each pair was presented five times, resulting in 15 trials per participant. Products were presented together onscreen for 3000ms, and each pair was separated by an intertrial interval of 1000 ms.

Rating phase. Participants evaluated the three target products on 21 traits and then evaluated them on their fat content.[10] As before, products were displayed at the top of the screen, with the rating scales below.

Pairing memory. As in Studies 2-4, participants were asked to indicate which source product each target product had been presented with during the pairing phase.

Source feature memory. The source products were presented onscreen without their fat content information along with a filler picture (a cookie pack not seen before). Participants were asked to indicate for each product whether it had previously been labelled as low, medium, or high-fat content or none of the three. This manipulation check was pre-registered and carried out to exclude participants who could not correctly remember the fat labelling. That said, a large number of participants failed to respond correctly to this manipulation (N = 48 responded incorrectly on at least one question). So we instead decided to test for the moderating impact of this variable on the halo and trait selectivity effects.

Finally, participants provided demographic information (age, gender, native language) and completed a series of exploratory questions (see Supplementary Materials).

Results

Two participants who failed to complete the study were excluded. We predicted a pairing-based halo effect that should vary as a function of trait halo relevance and pairing memory. Source feature memory did not moderate our main effects, so we removed this factor from our analyses. Mixed-models were the same as the ones in Study 5, except that we added memory as a fixed effect when testing for its moderation.

Fat content ratings (OLS regression). Participants judged the target products paired with high fat products as having more calories and fat than those paired with low fat products, $t(145) = 4.62$, $p < .001$, $dz = 0.38$, 95% CI [0.21; 0.55] ($M_{low} = 4.08$, $SD_{low} = 1.85$, $M_{med} = 4.73$, $SD_{med} = 1.49$, $M_{high} = 5.07$, $SD_{high} = 1.60$).

Pairing-based halo effect (mixed-model). On average, trait ratings did not differ significantly for target products paired with low fat products compared to those paired with high fat products (linear contrast C1), $t(37.64) = 0.68$, $p = .50$, ($M_{low} = 4.20$, $SD_{low} = 1.70$, $M_{med} = 4.12$, $SD_{med} = 1.69$, $M_{high} = 4.14$, $SD_{high} = 1.72$).

Trait selectivity: relevance- or valence-based? (mixed-model). An interaction emerged between the fat content linear contrast and trait relevance, $t(141.12) = 4.02$, $p < .001$, such that the pairing-based halo effect was significant and in the expected direction for traits high in halo relevance (+1*SD*), $t(138.61) = 4.00$, $p < .001$, and significant but in the opposite direction for traits low in halo relevance (-1*SD*), $t(141.84) = 2.46$, $p = .01$ (see Figure 7, *left panel*). This interaction effect was not moderated by trait valence, $t(162.10) = 0.35$, $p = .73$, and remained significant when trait valence was added to the model, $t(156.40) = 4.00$, $p < .001$. Finally, the pairing-based halo effect was not moderated by trait valence, $t(162.10) = 0.56$, $p = .58$ (see Figure 7, *right panel*).

Moderation by pairing memory (mixed-model). The pairing-based halo effect was not moderated by pairing memory ($N_{correct} = 70$, $N_{incorrect} = 76$), $t(180.15) = 0.56$, $p = .58$. The relevance-based trait selectivity, however, was moderated by pairing memory, $t(158.84) = 3.20$, $p = .002$,

9 Demographic information was absent for two participants.

10 This departs from Studies 2-4 where we measured the attractiveness of both source and target faces. In the present case, however, the objective metric of the source products' fat content provided by the instructions gives little space for ambiguity on this dimension so we decided not to include this measure for the source products.

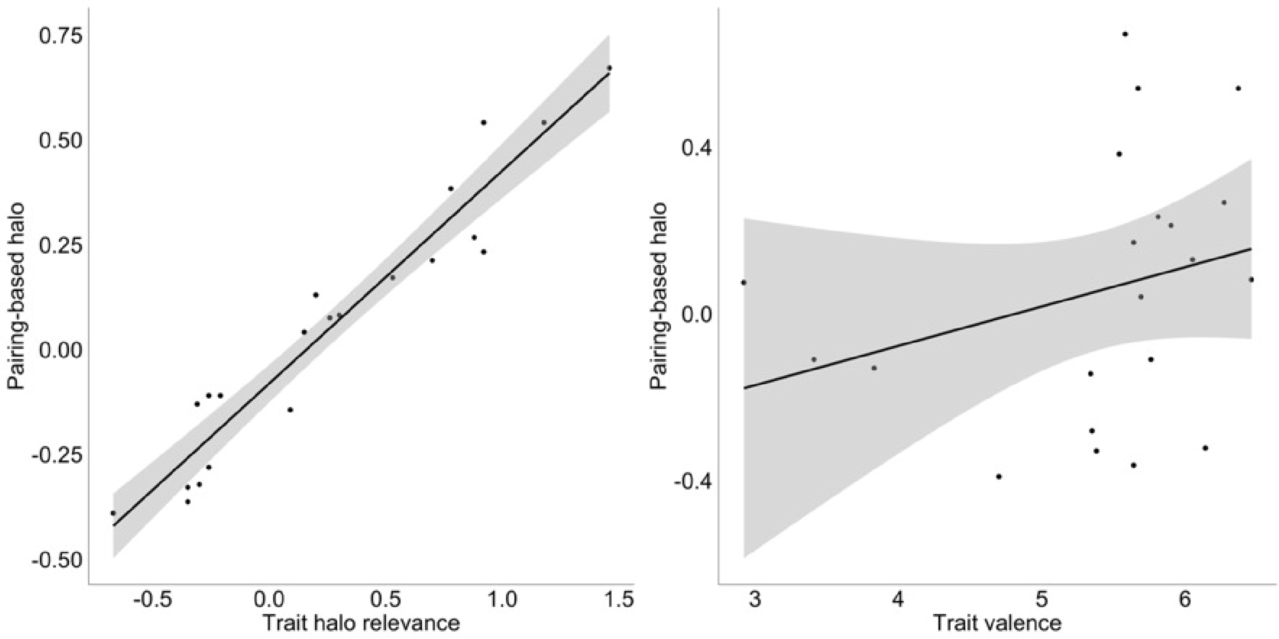

Figure 7. Pairing-based halo effect (Low-High scores) at the trait level, as a function of trait halo relevance (left panel) and trait valence (right panel)

Note. Grey areas represent the 95% confidence intervals. The pairing-based halo effect was positively related to both trait halo relevance (*left panel*). It was not significantly related to trait valence (*right panel*).

with larger trait selectivity occurring for participants with correct compared to incorrect pairing memory. This moderation was not significant for valenced-based trait selectivity, $t(8293.85) = 0.33$, $p = .74$. The overall interaction between the fat content linear contrast, trait halo relevance, trait valence, and pairing memory was non-significant, $t(8293.85) = 1.91$, $p = .06$.

Discussion

A pairing-based health halo effect emerged for traits high in relevance. The pairing-based halo effect was relevance-based, as predicted. Pairing memory did not influence the pairing-based halo effect, but it did influence its relevance-based trait selectivity. The trait selectivity makes it harder to detect changes in the average pairing-based as a function of pairing memory: because changes occur in both positive and negative directions, they tend to cancel each other out. When, however, the halo relevance of the traits is taken into account, the influence of the pairing memory clearly emerges: the pairing-based halo effect is more extreme (in a positive or a negative direction, as a function of the trait halo relevance) when the memory is correct than incorrect.

General Discussion

For most researchers, halo effects in social psychology and conditioning effects in learning psychology represent two distinct and unrelated phenomena. However, a recently proposed framework suggests that halo and conditioning effects may be described as instances of feature transformation effects (De Houwer et al., 2019). This framework also argues that there may be new empirical phenomena at the crossroads of halo and conditioning that are still waiting to be discovered. We systematically explored one such new phenomenon which we labelled the "pairing-based halo effect". This effect involves a situation with (a) different source and target features, as in halo studies, and (b) different source and target objects, as in conditioning studies. We tested if such an effect would emerge and whether it is influenced by the same moderators that impact halo (i.e., halo relevance of the traits) and conditioning effects (i.e., pairing memory and feature salience). We addressed these questions across six pre-registered studies and with two types of halo effects (i.e., attractiveness halo vs. health halo).

We first set out to replicate the classic halo effects for attractiveness (Study 1; e.g., Dion et al., 1972) and health (Study 5; e.g., Westcombe & Wardle, 1997). In both cases the effects varied as a function of the target feature (trait) being assessed and this variation was independent of the valence of the trait assessed. These studies also allowed us to generate continuous "trait halo relevance" scores (i.e., scores reflecting the average halo effect produced by each target trait) that we used in the following pairing-based studies to test whether the selectivity was similar (i.e., relevance-based).

In Studies 2-4 we set out to examine if a pairing-based halo effect could be established where attractiveness was the source feature. Source objects (faces) that varied in their attractiveness (low vs. medium vs. high) were paired with target objects (faces) that were medium in their attractiveness. In our first attempt (Study 2), the pairing-based halo effect varied across traits as a function of the valence of the trait (i.e., it was valence-based), not as a function of the halo relevance of the trait (i.e., it was not relevance-based). Relying on past EC and AC effects indicating that feature salience can play an important role when a object has multiple different features (e.g., Förderer & Unkelbach, 2014; Gast & Rothermund, 2011), we conducted a new study in which the source feature of attractiveness was made salient (Study 3). This time we found a pairing-based halo effect that was relevance-based. Both pairing-based halo effect and relevance-based trait selectivity were larger when participants had a correct (vs. incorrect) memory of the source and target objects pairs.

Study 4 directly tested the importance of the source feature salience in pairing-based halo effects. We compared a condition where attractiveness was salient to one where it was not. An overall pairing-based halo effect emerged that was larger when attractiveness was made salient. The nature of the trait selectivity varied as a function of the salience condition, such that it was valence-based in the no salience condition but relevance-based in the salience condition, replicating what we found in Studies 2 and 3. Again, pairing memory moderated the magnitude of both the pairing-based halo effect and its relevance-based trait selectivity.

Finally, Study 6 examined if similar findings would emerge in the context of a health halo effect. Source objects (cookie packs) varying in their fat content (low vs. medium vs. high) were paired with target objects (cookie packs) for which no fat content information was provided. The pairing-based halo effect strongly depended on trait halo relevance (and not on trait valence) and only emerged for traits for which standard health halo effects were also large (i.e., high trait halo relevance). Pairing memory did not influence the pairing-based effect but did influence the relevance-based trait selectivity.

Taken together, our studies showcase how effects traditionally assumed to be distinct can in fact be related in order to uncover novel empirical phenomena. Of note, our effects emerged for different types of halo (i.e., attractiveness and health), using various populations (English-speaking vs. Italian-speaking), and using mixed-models (i.e., analyses known to lead to more robust and generalisable results). The large sample sizes used in our experiments provide additional confidence in the reliability of our findings. In what follows we unpack the empirical and theoretical implications of our findings for both the impression formation (social psychology) and conditioning (learning psychology) literatures.

Empirical Contributions to the Impression Formation and Learning Literatures

Our results contribute to both the social psychology (impression formation) and learning psychology (conditioning) literatures. First, they highlight co-occurrence of stimuli as a pathway for establishing and changing our impressions of others (i.e., via the spatiotemporal contiguity between source and target objects; also see Moran et al., 2022). Specifically, we learned more about the context under which the manipulation of a source feature influence target traits of a target object. In our studies the impression formed about a source object (e.g., person) on the basis of a salient feature (e.g., attractiveness) seems to "radiate" out to a target object when the two were presented together in space and time. A unique feature of this approach to changing impressions is that it does not specify how the source and target objects are related. This departs from the "radiating beauty effect" that only emerges with an established assimilative relation between the source and target objects (e.g., two persons described as romantic partners; Sigall & Landy, 1973). The fact that we did find halo-like effects without specifying any relation might be because we repeated paired source and target object in a salient manner. It has been suggested that repeated salient pairings might function as a cue for assuming an assimilative relation between the paired objects (raising the possibility that conditioning effects emerge from assimilation processes; see De Houwer & Hughes, 2016).

Interestingly, a similar effect can be observed in the Spontaneous Trait Transference (STT) effect (similar to the "kill-the-messenger" effect; Manis et al., 1974) that relies on the pairing between a photographed person and a behavioural information about another person (Skowronski et al., 1998). In the STT effect, the source and the target objects are different (the source object is the actor of the behaviour and the target object the photographed person). The source feature (a behavioural statement such as "He never says thank you") and target features (a trait such as impolite) are in a different format but refer to the same feature (i.e., politeness). As in the pairing-based halo effect, the STT effect can emerge even without defining an assimilative relation between the actor and the photographed person (e.g., ostensible random pairing; Goren & Todorov, 2009; Skowronski et al., 1998). The (pairing-based) halo effect and the STT differ, however, with regard to the nature of the source feature (i.e., visual information versus behavioural statement).

Second, our results provide new insights into the trait selectivity of classic halo effects (see also Han & Laurent, 2023). We show for the first time that the trait selectivity of halo effects can go beyond valence while *statistically controlling* for it. This contradicts the old idea that observing one good feature of an object makes it generally more positive (e.g., "what is beautiful is good" in the case of the attractiveness halo; Dion et al., 1972). Earlier work already showed that the halo effect could involve both positive and negative effects (e.g., for the attractiveness halo; Bassili, 1981; Dermer & Thiel, 1975; and for the health halo; Lee et al., 2013; Orquin & Scholderer, 2015; Westcombe & Wardle, 1997). Critically, previous studies did not test whether (a) the traits for which the halo effect is the strongest are also the most positive (i.e., whether trait relevance and valence are confounded) and, more importantly, (b) whether the observed trait selectivity is still present after controlling for trait valence. In Studies 1 and 5, mixed-models enabled us to rule out the role of valence by opposing within the same model the influence of trait relevance and valence. We considered target features as a random effect and thus modelled variables inherent to this factor (relevance and valence). A positive side effect is that the observed effects can be generalised to other participants and other target features with similar properties (i.e., the pairing-based halo effect should be observed for different relevant traits; Judd et al., 2012).

Overall, our results are consistent with theoretical accounts stating that the halo effect reflects the shared structure of beliefs toward certain categories of stimuli/traits (e.g., implicit theories about attractive people/attractiveness; e.g., see Forgas & Laham, 2016). For instance, the trait selectivity of the attractiveness halo effect might reflect the stereotype about (un)attractive people that are seen as more (less) social, glamorous, and vain but have less (more) integrity. We will return to this idea in the next section.

The third contribution of our work relates to the target features generally used in the radiating beauty effect (e.g., Sigall & Landy, 1973). Earlier studies on the radiating beauty effect used multiple target features (e.g., self-confident, likable, talented), but those earlier studies did not include a comparison of the effects for different target features. We are the first to examine trait selectivity effects with different source and target objects by using target features used in the classic halo effect (e.g., also negative traits such as vanity; Bassili, 1981). As noted above, the radiating beauty phenomenon differs in several aspects from the pairing-based halo effect. Nevertheless, our work suggests that relevance-based trait selectivity could also be a property of the radiating beauty effect.

Finally, our results also contribute to the conditioning literature. As explained in the Introduction, the overwhelming majority of EC and AC studies have been limited to identical source and target features (e.g., valence).[11] One exception can be noted in work of Kim et al. (1996) where the authors demonstrated an AC effect where the source object (e.g., speed from a race car) influenced how a target object (e.g., pizza brand) was perceived on a series of attributes. They observed that only related attributes (e.g., delivered hot and reliable) were influenced but not unrelated attributes. However, authors did not test whether the AC effect was *larger* for related than unrelated attributes (i.e., whether the type of attribute moderated the AC effect) or whether the observed effects were due to mere valence (e.g., whether the influence of speed only applied more extreme positive attributes than less extreme ones). Hence, our work is first in showing that pairing effects selectively apply on different source and target features while controlling for trait valence as well. It also underlies how surprising it is to see that target features that differ from the source features are typically not assessed in the AC literature. For instance, in an AC study using athletic as a source and target feature (Förderer & Unkelbach, 2011), participants could have been asked to rate various target features that stereotypically relate to athleticism (e.g., leadership, dominance). More systematic testing of changes on different target features – especially those stereotypically related to the source feature – would allow researchers to assess selectivity in these effects and the sources of this selectivity (e.g., prior beliefs, trait valence). In this regard, the statistical method that we adopted (i.e., opposing within the same statistical model the trait values of halo vs. valence) is well suited to address the question of the source of the selectivity.

Theoretical Implications

Our findings also shed some light on the mental mechanisms that may mediate pairing-based halo effects. They allow us to assess whether these effects can be accounted for by existing impression formation and/or conditioning theories.

As outlined in the introduction, accounts of halo effects often refer to the idea of stereotyping or implicit theories of personality to explain the selective and consistent influence of one feature (e.g., physical attractiveness) on multiple other features (e.g., social competence, vanity). These accounts seem perfectly suited to explain trait selectivity of halo effects: when a source feature (or a category of persons possessing this feature) and a target feature are thought to be conceptually related, activating one triggers the other (Schneider, 1973). Interestingly, this account would also explain the role of source feature salience: the salience of a stimulus feature (e.g., attractiveness) or category (e.g., attractive people) is likely to determine the stereotype or implicit personality theory that become activated (e.g., the stereotype of attractive people) and thus the observed changes in ratings (Crawford et al., 2002; Fiske & Neuberg, 1990). However, stereotypes and implicit personality theories are typically about a specific person whereas pairing-based effects refer to assumptions about a target object different from the source object.

One possible solution could be offered by invoking the general principles of the similarity heuristic: when two persons or two objects are presented together in space and time, participants may assume that they are similar in other respects (e.g., De Houwer & Hughes, 2016; Hughes et al., 2020; see also Carlston & Skowronski, 2005). This heuristic would explain the relevance-based trait selectivity observed in our pairing-based halo studies: an individual paired with an attractive person may be perceived as possessing the same traits as the attractive person (e.g., social competence). Finally, this reasoning would also be consistent with

11 Interestingly, this limitation also applies for the effect of attractiveness in social context which is more typical of impression formation (Melamed & Moss, 1975). Research on attractiveness context effects shows that flanking a face with two other faces that are high (vs. low) on attractiveness shifts the perceived attractiveness of the target face (e.g., Geiselman et al., 1984; Rodway et al., 2013). This phenomenon is similar to AC in the sense that it focuses on the same source and target feature, however, the two differ on several other aspects. For instance, the pairing generally involves more than two stimuli (Geiselman et al., 1984), and the ratings of the target face sometimes take place during the pairing procedure and in the presence of the flankers (Rodway et al., 2013).

the influence of pairing memory: only when participants remember what stimuli go together can they rely on this similarity heuristic.

This latter interpretation also relates to current explanations of the "radiating beauty" effect, except that the latter relies on more complex heuristics. For instance, the partner of an attractive person is thought to possess more positive traits because the association with the attractive person signals that the partner may have desirable qualities for mating choice (heuristic based on indirect cues such as "he must have something more than others"; Sigall & Landy, 1973). In this case, it is not only the source feature and the assimilative relation between the objects *per se* that leads to the attributions about the target object, but also additional indirect cues (e.g., the "romantic partners" relation between the source and target objects). Interestingly, the interplay between multiple cues (e.g., gender of the source object) could result in different attributions and thus different trait selectivity than the one found in the present work. For instance, Kocoglu and Mithani (2020) showed that the more the (female) partner of a male individual was attractive, the more he was perceived as a leader. However, the reversed effect was observed when the target was a woman – the more attractive her (male) partner, the less dominant and competent she was perceived. These results raise the possibility of "relational stereotypes", that is, stereotypes based on how an object is related to other objects. For instance, in the study of Kocoglu and Mithani (2020), effects might reflect the stereotype about male partners of attractive females rather than the mere stereotype about attractive people.

It is also informative to relate our work to theoretical accounts of conditioning effects. Conditioning effects are typically explained using associative (via the formation of associations; Rescorla & Wagner, 1972) or propositional accounts (via the formation of propositional representations; De Houwer, 2009, 2018; Mitchell et al., 2009). Both approaches would predict the moderating role of the source feature salience on the pairing-based halo effect. Following the propositional approach, for instance, making the source feature salient would facilitate inferences involving this feature. Regarding the effect of pairing memory, the exact mechanisms (e.g., lack of attention and effort) that underlie this effect are still under investigation, despite many years of research in EC (e.g., Blask et al., 2017; Field & Moore, 2005). Nevertheless, the moderating role of pairing memory has been a crucial factor in developing EC theories as the associative and propositional theoretical approaches formulate different predictions. The importance of the pairing memory fits well with the idea that the formation of propositions about pairings requires awareness of those pairings (De Houwer, 2009). According to most associative accounts, however, the formation of associations in memory does not depend on the awareness of the contingencies (Corneille & Stahl, 2018; De Houwer, 2011). Both propositional and associative accounts could, in principle, accommodate that a source feature of a source object influences other target features of a target object. For instance, one could argue that responses elicited by the source object (e.g., higher social competence perception elicited by attractive people) become associated with the target object as a result of repeated pairings. From a propositional perspective, inferences can be based on propositions about the co-occurrence of the source and target objects and the source feature. For instance, the propositions emerging from the pairings and the source feature (e.g., "this attractive person goes together with this other person") mixed with propositions about previous knowledge (e.g., "what goes together is usually similar") would lead to inferences about the target object (e.g., "the other person is likely to be socially competent"; see also De Houwer & Hughes, 2016; Moran et al., 2022). Interestingly, this propositional account is highly similar to attribution accounts in the impression formation literature.

However, neither associative nor propositional accounts provide a straightforward explanation for the relevance-based trait selectivity of the pairing-based halo effect. Any cognitive account needs to refer to conceptual knowledge on how features relate to each other. Stereotypes and implicit theories of personality specify such knowledge but have never been considered in conditioning theories. Yet, conceptual knowledge about the relation between the source feature and other target features or even between the source and target objects could play a crucial role in known learning effects. Consider, for instance, the possibility that people have pre-existing stereotypes about persons who are often seen together with other liked people. If these stereotypes specify that persons who co-occur with liked people are themselves positive, then EC effects might be the product of applying these pre-existing stereotypes to a novel person that co-occurs with liked other people – in much the same way as attractiveness halo effects would result from applying stereotypes about attractive people to a novel attractive person. These ideas fit well with the more general idea that organisms deploy all kinds of pre-existing knowledge when taking part in a learning experiment. Therefore, the learning effects observed in experiments might be a joined product of the events taking place during the experiment and the knowledge that organisms have acquired pre-experimentally.

Our theoretical considerations show that accounts of impression formation and learning can contribute to the explanation of the pairing-based halo effect. They thereby highlight the merits of breaking down the unnecessary barriers between these two research fields. Concepts not commonly used in one theoretical domain (e.g., stereotypes) could be valuable in another (e.g., conditioning accounts). These theoretical interactions are made possible by the conceptual feature transformation framework that allows using a common language to refer to procedures and effects in impression formation research and learning research.

Limitations and Future Directions

Like any paper, ours is subject to several limitations. First, we only found a significant average pairing-based halo effect in Study 4. This finding makes sense, however, if we take into account the moderators of the effect. The most important moderator is the type of target trait that is assessed, as illustrated by the relevance-based trait selectivity. Studies 3, 4, and 6 revealed a pairing-based halo effect for traits high in halo relevance. Pairing-based halo effects also emerged reliably when the source feature was salient, and participants had a correct memory of the pairing. Therefore, we can conclude that the pairing-based halo effect is robust provided that boundary conditions are met.

Second, although this was not the main purpose of our work, the causal chain via which the source feature influenced judgements about the target features is not clear yet. At this stage, we see two potential intermediary mechanisms. A first possibility is that the manipulation of the source feature on the source object (e.g., attractiveness) influenced the perception of the target object on the same feature (e.g., attractiveness), which in turn influenced the ratings of the target object on the target features (e.g., increase in social competence ratings). Exploratory analyses of Study 4 (see Supplementary Materials) suggest that this could be the case while also showing that this mediation process was moderated by the source feature salience (i.e., the effect of the source object's attractiveness on the target object's attractiveness was larger when the source feature was salient). This would be consistent with an assimilation effect between the two faces on the source feature (i.e., consistent with an AC effect or a similarity heuristic) followed by a halo effect on the target object.

Importantly, however, a mediation was observed only in Study 4. We also did not consistently observe changes in perception of the target object on the manipulated source feature – we did observe this change in Study 6 when using food products but not in the studies with faces. We believe that the lack of effect of source attractiveness on judgements about the attractiveness of the target face is due to the use of face stimuli. Specifically, it might be difficult to influence how attractive a face is perceived because faces have visual features that constrain judgements of attractiveness. Notably, this departs from AC research in which CSs are made ambiguous on the manipulated attribute (e.g., face pictures when manipulating athleticism; Förderer & Unkelbach, 2011). On the contrary, using personality traits/outcomes, that is, information less grounded in visual features (but highly relevant for the manipulated source feature), leads to strong pairing-based effects.

Another possibility is that the manipulation of the source feature on the source object (e.g., attractiveness) influenced how the source object is perceived on the features used as targets (e.g., increase in social competence perception), which in turn influenced the ratings of the target object on the same features (e.g., increase in social competence ratings). This would be consistent with a halo effect on the source object followed by an assimilation effect between the two faces on the target features (i.e., consistent with an AC effect or a similarity heuristic). Current data, however, do not allow us to assess this possibility because we did not collect target feature ratings on the source objects in pairing studies.

Clearly there are several possible mechanistic accounts for the findings reported here. We offer our findings as a springboard for discussions about those mechanisms and a starting point for future work testing competing ideas. Researchers could investigate potential causal chains with a full experimental design (e.g., interruption of the causal process with a contextual variable; Jacoby & Sassenberg, 2011). A mediation effect via changes in perceived attractiveness of the target face vs. via changes in perceived target features of the source face would inform on which of the first or second possibilities is the most probable, respectively.

Third, the impact of source feature salience and memory of the pairings raises the question of the potential role of demand awareness (perceived wishes of the experimenter) and demand compliance (compliance to the perceived wished of the experimenter) in pairing-based halo effects. For instance, because participants were instructed to focus on the source feature and on the pairings, they may have felt pressured into treating the target object as similar to the source object in terms of the source feature (e.g., as highly attractive if the source object was a highly attractive face). We also see at least four reasons why demand compliance cannot explain our results. First, if participants perceived a demand from the experimenter, such demand was probably related to how they should evaluate the target object on attractiveness/fat content (e.g., "rate the attractiveness of the target in the same way as the attractiveness of the source"). However, we did not observe this effect. Second, whereas demand compliance is possible only if participants correctly memorize the pairings, exploratory analyses (see Table S6 in Supplementary Material) showed that changes in perceived attractiveness (or fat content) of the target object were moderated by the memory of the pairings only in Study 6. Third, we see no reasons why demand effects would explain the observed trait selectivity effect, that is, the observation that the source feature influenced ratings more strongly for some than for other target features (e.g., larger effect on social competence than integrity). Even if participants would have ideas about the exact pattern of selectivity that the researchers expected, the perceived demand for selectivity would probably be much weaker than the perceived demand for the manipulated source feature (e.g., attractiveness). The combined facts that no effects were observed for the manipulated source feature but trait selectivity effects were observed renders it highly unlikely that demand compliance was a major factor in our studies. Finally, when we performed additional analyses after excluding participants categorized as demand aware (i.e., participants who mentioned any potential effect of pairings in the open-ended question at the end of the experiment) or reporting demand compliance, we found that results remained unchanged (see Table S7 in Supplementary Material).

Finally, the effects of the moderators we examined do not allow for unambiguous conclusions about the processes underlying the pairing-based halo effect. The moderating role of the pairing memory is in line with a propositional approach but contradicts most associative accounts. Yet, further work is still needed. Interestingly, the two main differences between the propositional and associative accounts concern the role of relational information (i.e., how stimuli are related) and truth value (i.e., whether the encoded information is true or false; De Houwer, 2009, 2018). Future research could thus explore whether information about the relation between the source and the target objects moderates the pairing-based halo effect. In line with this idea, the radiating beauty effect is strongly moderated by the relational information between the source and target objects (e.g., in couple vs. unrelated; Sigall & Landy, 1973). In the same vein, relational information (e.g., friends or enemies; Fiedler & Unkelbach, 2011) has proven to be an important moderator of EC effects. Moderation of the pairing-based halo effect by relational information would favour a propositional account of these effects.

Conclusion

Drawing on a new conceptual framework we introduce and systematically test a novel phenomenon at the crossroads between the halo and conditioning literatures. This phenomenon, which we referred to as the pairing-based halo effect, is sensitive to several key moderators of impression formation and conditioning effects, further emphasizing its hybrid nature. Although more work is needed on the causal chain or specific processes underlying this effect, our findings raise the possibility that common principles may govern impression formation and conditioning effects.

Competing Interests

Authors report having no conflicts of interest with the publication of the present work.

Contributions

Contributed to conception and design: MR, JDH, JR, and MP

Contributed to acquisition of data: MR and JR

Contributed to analysis and interpretation of data: MR and JR

Drafted and/or revised the article: MR, JDH, JR, SH, and MP

Approved the submitted version for publication: MR, JDH, JR, SH, and MP

Funding

This research was made possible by the Methusalem grant BOF16/MET_V/002 of Ghent University to Jan De Houwer.

Data Accessibility Statement

The pre-registration files, deviations from the pre-registrations document, materials (including stimuli and JsPsych or Inquisit code), data, and analytic (R) scripts for all experiments are made publicly available at https://osf.io/5dftn/?view_only=fc2c3f2e2e0e402d93d7d2b81a57a284.

Submitted: April 11, 2023 PDT, Accepted: June 15, 2023 PDT

References

Andrews, J. C., Netemeyer, R. G., & Burton, S. (1998). Consumer generalization of nutrient content claims in advertising. *Journal of Marketing*, *62*(4), 62–75. https://doi.org/10.1177/002224299806200405

Andrews, J. C., Netemeyer, R. G., & Burton, S. (2009). The nutrition elite: Do only the highest levels of caloric knowledge, obesity knowledge, and motivation matter in processing nutrition ad claims and disclosures? *Journal of Public Policy & Marketing*, *28*(1), 41–55. https://doi.org/10.1509/jppm.28.1.41

Ashmore, R. D. (1981). Sexstereotypes and implicit personality theory. In D. L. Hamilton (Ed.), *Cognitive processes in stereotyping and intergroup behavior* (pp. 37–81). Erlbaum.

Bainbridge, W. A., Isola, P., & Oliva, A. (2013). The intrinsic memorability of face photographs. *Journal of Experimental Psychology: General*, *142*(4), 1323–1334. https://doi.org/10.1037/a0033872

Bassili, J. N. (1981). The attractiveness stereotype: Goodness or glamour? *Basic and Applied Social Psychology*, *2*(4), 235–252. https://doi.org/10.1207/s15324834basp0204_1

Blask, K., Walther, E., & Frings, C. (2017). When congruence breeds preference: The influence of selective attention processes on evaluative conditioning. *Cognition and Emotion*, *31*(6), 1127–1139. https://doi.org/10.1080/02699931.2016.1197100

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*(6), 884–898. https://doi.org/10.1037/0022-3514.89.6.884

Corneille, O., & Stahl, C. (2018). Associative attitude learning: A closer look at evidence and how it relates to attitude models. *Personality and Social Psychology Review*, *23*(2), 161–189. https://doi.org/10.1177/1088868318763261

Crawford, M. T., Sherman, S. J., & Hamilton, D. L. (2002). Perceived entitativity, stereotype formation, and the interchangeability of group members. *Journal of Personality and Social Psychology*, *83*(5), 1076–1094. https://doi.org/10.1037/0022-3514.83.5.1076

De Houwer, J. (2007). A conceptual and theoretical analysis of evaluative conditioning. *The Spanish Journal of Psychology*, *10*(2), 230–241. https://doi.org/10.1017/s1138741600006491

De Houwer, J. (2009). The propositional approach to associative learning as an alternative for association formation models. *Learning & Behavior*, *37*(1), 1–20. https://doi.org/10.3758/lb.37.1.1

De Houwer, J. (2011). Evaluative Conditioning: A Review of Procedure Knowledge and Mental Process Theories. In T. R. Schachtman & S. Reilly (Eds.), *Associative Learning and Conditioning Theory: Human and Non-Human Applications* (pp. 399–416). Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199735969.003.0130

De Houwer, J. (2018). Propositional models of evaluative conditioning. *Social Psychological Bulletin*, *13*(3), 1–21. https://doi.org/10.5964/spb.v13i3.28046

De Houwer, J., & Hughes, S. (2016). Evaluative conditioning as a symbolic phenomenon: On the relation between evaluative conditioning, evaluative conditioning via instructions, and persuasion. *Social Cognition*, *34*(5), 480–494. https://doi.org/10.1521/soco.2016.34.5.480

De Houwer, J., & Hughes, S. (2020). *The psychology of learning: An introduction from a functional-cognitive perspective*. MIT Press.

De Houwer, J., Richetin, J., Hughes, S., & Perugini, M. (2019). On the Assumptions That We Make About the World Around Us: A Conceptual Framework for Feature Transformation Effects. *Collabra: Psychology*, *5*(1). https://doi.org/10.1525/collabra.229

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. https://doi.org/10.1037/0033-2909.127.6.853

de Leeuw, J. R. (2015). jsPsych: A JavaScript library for creating behavioral experiments in a web browser. *Behavior Research Methods*, *47*(1), 1–12. https://doi.org/10.3758/s13428-014-0458-y

Dermer, M., & Thiel, D. L. (1975). When beauty may fail. *Journal of Personality and Social Psychology*, *31*(6), 1168–1176. https://doi.org/10.1037/h0077085

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology*, *24*(3), 285–290. https://doi.org/10.1037/h0033731

Eagly, A. H., Ashmore, R. D., Makhijani, M. G., & Longo, L. C. (1991). What is beautiful is good, but…: A meta-analytic review of research on the physical attractiveness stereotype. *Psychological Bulletin*, *110*(1), 109–128. https://doi.org/10.1037/0033-2909.110.1.109

Ebneter, D. S., Latner, J. D., & Nigg, C. R. (2013). Is less always more? The effects of low-fat labeling and caloric information on food intake, calorie estimates, taste preference, and health attributions. *Appetite*, *68*, 92–97. https://doi.org/10.1016/j.appet.2013.04.023

Efran, M. G. (1974). The effect of physical appearance on the judgment of guilt, interpersonal attraction, and severity of recommended punishment in a simulated jury task. *Journal of Research in Personality*, *8*(1), 45–54. https://doi.org/10.1016/0092-6566(74)90044-0

Fiedler, K., & Unkelbach, C. (2011). Evaluative conditioning depends on higher order encoding processes. *Cognition & Emotion*, *25*(4), 639–656. https://doi.org/10.1080/02699931.2010.513497

Field, A. P., & Moore, A. C. (2005). Dissociating the effects of attention and contingency awareness on evaluative conditioning effects in the visual paradigm. *Cognition & Emotion*, *19*(2), 217–243. https://doi.org/10.1080/02699930441000292

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: influences of information and motivation on attention and interpretation. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 23, pp. 1–74). Academic Press. https://doi.org/10.1016/s0065-2601(08)60317-2

Förderer, S., & Unkelbach, C. (2011). Beyond evaluative conditioning! Evidence for transfer of non-evaluative attributes. *Social Psychological and Personality Science*, *2*(5), 479–486. https://doi.org/10.1177/1948550611398413

Förderer, S., & Unkelbach, C. (2014). The moderating role of attribute accessibility in conditioning multiple specific attributes. *European Journal of Social Psychology*, *44*(1), 69–81. https://doi.org/10.1002/ejsp.1994

Förderer, S., & Unkelbach, C. (2015). Attribute conditioning: Changing attribute-assessments through mere pairings. *Quarterly Journal of Experimental Psychology*, *68*(1), 144–164. https://doi.org/10.1080/17470218.2014.939667

Forgas, J. P., & Laham, S. M. (2016). Halo effects. In *Cognitive Illusions* (pp. 286–300). Psychology Press. https://doi.org/10.4324/9781315696935-22

Forwood, S. E., Walker, A. D., Hollands, G. J., & Marteau, T. M. (2013). Choosing between an apple and a chocolate bar: The impact of health and taste labels. *PloS One*, *8*(10), e77500. https://doi.org/10.1371/journal.pone.0077500

Gast, A., & Rothermund, K. (2011). What you see is what will change: Evaluative conditioning effects depend on a focus on valence. *Cognition and Emotion*, *25*(1), 89–110. https://doi.org/10.1080/02699931003696380

Geiselman, R. E., Haight, N. A., & Kimata, L. G. (1984). Context effects on the perceived physical attractiveness of faces. *Journal of Experimental Social Psychology*, *20*(5), 409–424. https://doi.org/10.1016/0022-1031(84)90035-0

Gleibs, I. H. (2017). Are all "research fields" equal? Rethinking practice for the use of data from crowdsourcing market places. *Behavior Research Methods*, *49*(4), 1333–1342. https://doi.org/10.3758/s13428-016-0789-y

Goren, A., & Todorov, A. (2009). Two faces are better than one: Eliminating false trait associations with faces. *Social Cognition*, *27*(2), 222–248. https://doi.org/10.1521/soco.2009.27.2.222

Han, D. E., & Laurent, S. M. (2023). Beautiful seems good, but perhaps not in every way: Linking attractiveness to moral evaluation through perceived vanity. *Journal of Personality and Social Psychology*, *124*(2), 264–286. https://doi.org/10.1037/pspa0000317

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. https://doi.org/10.1037/a0018916

Hughes, S., De Houwer, J., Mattavelli, S., & Hussey, I. (2020). The shared features principle: If two objects share a feature, people assume those objects also share other features. *Journal of Experimental Psychology: General*, *149*(12), 2264–2288. https://doi.org/10.1037/xge0000777

Irwin, J. R., & McClelland, G. H. (2003). Negative consequences of dichotomizing continuous predictor variables. *Journal of Marketing Research*, *40*(3), 366–371. https://doi.org/10.1509/jmkr.40.3.366.19237

Jacoby, J., & Sassenberg, K. (2011). Interactions do not only tell us when, but can also tell us how: Testing process hypotheses by interaction. *European Journal of Social Psychology*, *41*(2), 180–190. https://doi.org/10.1002/ejsp.762

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology*, *103*(1), 54–69. https://doi.org/10.1037/a0028347

Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, *68*(1), 601–625. https://doi.org/10.1146/annurev-psych-122414-033702

Kocoglu, I., & Mithani, M. A. (2020). Does an attractive partner make you a better leader? Only if you are a male! *The Leadership Quarterly*, *31*(2), 101339. https://doi.org/10.1016/j.leaqua.2019.101339

Kuznetsova, A., Brockhoff, P. B., & Christensen, R. H. B. (2017). **lmerTest** Package: Tests in Linear Mixed Effects Models. *Journal of Statistical Software*, *82*(13), 1–26. https://doi.org/10.18637/jss.v082.i13

Landy, D., & Sigall, H. (1974). Beauty is talent: Task evaluation as a function of the performer's physical attractiveness. *Journal of Personality and Social Psychology*, *29*(3), 299–304. https://doi.org/10.1037/h0036018

Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin*, *126*(3), 390–423. https://doi.org/10.1037/0033-2909.126.3.390

Lee, W.-C. J., Shimizu, M., Kniffin, K. M., & Wansink, B. (2013). You taste what you see: Do organic labels bias taste perceptions? *Food Quality and Preference*, *29*(1), 33–39. https://doi.org/10.1016/j.foodqual.2013.01.010

Manis, M., Cornell, S. D., & Moore, J. C. (1974). Transmission of attitude relevant information through a communication chain. *Journal of Personality and Social Psychology*, *30*(1), 81–94. https://doi.org/10.1037/h0036639

McClelland, G. H., Lynch, J. G., Jr., Irwin, J. R., Spiller, S. A., & Fitzsimons, G. J. (2015). Median splits, Type II errors, and false–positive consumer psychology: Don't fight the power. *Journal of Consumer Psychology*, *25*(4), 679–689. https://doi.org/10.1016/j.jcps.2015.05.006

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, *17*(3), 437–455. https://doi.org/10.1037/a0028085

Melamed, L., & Moss, M. K. (1975). The effect of context on ratings of attractiveness of photographs. *The Journal of Psychology*, *90*(1), 129–136. https://doi.org/10.1080/00223980.1975.9923935

Mitchell, C. J., De Houwer, J., & Lovibond, P. F. (2009). The propositional nature of human associative learning. *Behavioral and Brain Sciences*, *32*(2), 183–198. https://doi.org/10.1017/s0140525x09000855

Moran, T., Hughes, S., Van Dessel, P., & De Houwer, J. (2022). The Role of Trait Inferences in Evaluative Conditioning. *Collabra: Psychology*, *8*(1). https://doi.org/10.1525/collabra.31738

Nisbett, R. E., & Wilson, T. D. (1977). The halo effect: Evidence for unconscious alteration of judgments. *Journal of Personality and Social Psychology*, *35*(4), 250–256. https://doi.org/10.1037/0022-3514.35.4.250

Olson, M. A., Kendrick, R. V., & Fazio, R. H. (2009). Implicit learning of evaluative vs. non-evaluative covariations: The role of dimension accessibility. *Journal of Experimental Social Psychology*, *45*(2), 398–403. https://doi.org/10.1016/j.jesp.2008.10.007

Oostenbach, L. H., Slits, E., Robinson, E., & Sacks, G. (2019). Systematic review of the impact of nutrition claims related to fat, sugar and energy content on food choices and energy intake. *BMC Public Health*, *19*(1), 1–11. https://doi.org/10.1186/s12889-019-7622-3

Orquin, J. L., & Scholderer, J. (2015). Consumer judgments of explicit and implied health claims on foods: Misguided but not misled. *Food Policy*, *51*, 144–157. https://doi.org/10.1016/j.foodpol.2015.01.001

Peer, E., Brandimarte, L., Samat, S., & Acquisti, A. (2017). Beyond the Turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology*, *70*, 153–163. https://doi.org/10.1016/j.jesp.2017.01.006

Peer, E., Vosgerau, J., & Acquisti, A. (2014). Reputation as a sufficient condition for data quality on Amazon Mechanical Turk. *Behavior Research Methods*, *46*(4), 1023–1031. https://doi.org/10.3758/s13428-013-0434-y

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. *Classical Conditioning II: Current Research and Theory*, *2*, 64–99.

Rodway, P., Schepman, A., & Lambert, J. (2013). The influence of position and context on facial attractiveness. *Acta Psychologica*, *144*(3), 522–529. https://doi.org/10.1016/j.actpsy.2013.09.004

Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*(4), 283–294. https://doi.org/10.1037/h0026086

RStudio Team. (2021). *RStudio: Integrated Development Environment for R*. RStudio. http://www.rstudio.com/

Schneider, D. J. (1973). Implicit personality theory: A review. *Psychological Bulletin*, *79*(5), 294–309. https://doi.org/10.1037/h0034496

Sigall, H., & Landy, D. (1973). Radiating beauty: Effects of having a physically attractive partner on person perception. *Journal of Personality and Social Psychology*, *28*(2), 218–224. https://doi.org/10.1037/h0035740

Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of the crime on juridic judgment. *Journal of Personality and Social Psychology*, *31*(3), 410–414. https://doi.org/10.1037/h0076472

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), 837–848. https://doi.org/10.1037/0022-3514.74.4.837

Stahl, C., & Heycke, T. (2016). Evaluative conditioning with simultaneous and sequential pairings under incidental and intentional learning conditions. *Social Cognition*, *34*(5), 382–412. https://doi.org/10.1521/soco.2016.34.5.382

Stahl, C., & Unkelbach, C. (2009). Evaluative learning with single versus multiple unconditioned stimuli: The role of contingency awareness. *Journal of Experimental Psychology: Animal Behavior Processes*, *35*(2), 286–291. https://doi.org/10.1037/a0013255

Vanaelst, J., Spruyt, A., Everaert, T., & De Houwer, J. (2017). Extinction of likes and dislikes: effects of feature-specific attention allocation. *Cognition and Emotion*, *31*(8), 1595–1609. https://doi.org/10.1080/02699931.2016.1250724

Walther, J. B., Van Der Heide, B., Kim, S.-Y., Westerman, D., & Tong, S. T. (2008). The role of friends' appearance and behavior on evaluations of individuals on Facebook: Are we known by the company we keep? *Human Communication Research*, *34*(1), 28–49. https://doi.org/10.1111/j.1468-2958.2007.00312.x

Wansink, B., & Chandon, P. (2006). Can "low-fat" nutrition labels lead to obesity? *Journal of Marketing Research*, *43*(4), 605–617. https://doi.org/10.1509/jmkr.43.4.605

Westcombe, A., & Wardle, J. (1997). Influence of relative fat content information on responses to three foods. *Appetite*, *28*(1), 49–62. https://doi.org/10.1006/appe.1996.0066

Westfall, J., Kenny, D. A., & Judd, C. M. (2014). Statistical power and optimal design in experiments in which samples of participants respond to samples of stimuli. *Journal of Experimental Psychology: General*, *143*(5), 2020–2045. https://doi.org/10.1037/xge0000014

Supplementary Materials

Peer Review History

Download: https://collabra.scholasticahq.com/article/84560-from-halo-to-conditioning-and-back-again-exploring-the-links-between-impression-formation-and-learning/attachment/173691.docx?auth_token=ihzTmW_cUwQ4Lv9w70Nu

Supplemental Materials

Download: https://collabra.scholasticahq.com/article/84560-from-halo-to-conditioning-and-back-again-exploring-the-links-between-impression-formation-and-learning/attachment/173692.docx?auth_token=ihzTmW_cUwQ4Lv9w70Nu