The Influence of Social Information and Presentation Interface on Aesthetic Evaluations

Yoko Urano[1], Raja Marjieh[2], Thomas L. Griffiths[2, 3], and Nori Jacoby[1]
[1]Computational Auditory Perception Group, Max Planck Institute for Empirical Aesthetics, Frankfurt am Main, Germany
[2]Department of Psychology, Princeton University, New Jersey, USA
[3]Department of Computer Science, Princeton University, New Jersey, USA

Abstract

We make aesthetic judgments on a daily basis. While we think of these judgments as highly personal, they are often shaped by social context. This poses a computational problem: how do we combine social information and our individual judgments to produce a single evaluation? In this study, we examine social influence on aesthetic evaluations in online transmission chain experiments. We test not only the effect of social information, but also variation in effect depending on how information is presented–echoing the variety of interfaces we encounter in naturalistic cases. We find that social information significantly affects ratings across interfaces. Moreover, people tend to rely more heavily on their own judgment than on social information, compared to an ideally noise-reducing model for combining multiple signals. These results offer detailed insight into the formation of aesthetic judgment and suggest the need for extended investigation into social influence on subjective judgments more broadly.

Keywords: social cognition; aesthetics; computational modeling; interface; psychology

Introduction

Whether choosing a song to listen to or a shirt to wear, we make dozens of aesthetic judgments every day. While we may think of these as a reflection of our unique personal tastes, our frequent exposure to others' aesthetic judgments—through ads, social media, or personal conversation—raises the question: to what extent are our aesthetic judgments shaped by the judgments of others? In this study, we investigate the contribution of social information towards aesthetic evaluations by gathering ratings of visual appeal for artworks in social and asocial contexts. We use multiple systematically manipulated interfaces for showing information in our experiments to derive a generalizable understanding of social influence that does not rely on a specific interface. Finally, we use computational modelling to identify the relative contribution of social information and prior individual judgment.

We know that people incorporate social information into their own judgments and decisions, both in objective tasks (e.g. Asch, 1951) and in subjective contexts (e.g. Verpooten & Dewitte, 2017). A large body of work has described the propagation of social information in objective tasks, including how and why these dynamics occur (Kendal et al., 2018; Soll & Larrick, 2009; Henrich & Gil-White, 2001); however, the dynamics of social influence in subjective tasks is less explored. A main goal of our study, therefore, is to address this through large scale empirical data.

In gathering this data, we also addressed the potential role of how we choose to present information to participants. To begin to understand the effects of interface features in social influence, we focus specifically on the absolute or relative nature of the information presented by the interface. While relative information strongly influences decision-making, such as in the case of positive and negative contrast (McNamara, Fawcett, & Houston, 2013), a direct comparison of absolute and relative information in the social comparison context show that people may rely more heavily on absolute information (Moore & Klein, 2008). Evidently, both absolute and relative information can play a significant role in judgment formation, and their effects depend on the context at hand. We therefore aim to clarify how these forms of information might affect social influence. To do so, we manipulate the absolute or relative nature of information through through our presentation of social information and stimuli. If absolute and relative information have significantly different effects on social influence in aesthetic judgments, we would expect to observe different dynamics across the interfaces we create through these manipulations.

The third component of this study is modelling the computational mechanism behind aesthetic judgments in social contexts. The aim is to derive a concrete idea of the contribution of social information, which allows for detailed comparisons between interfaces and conditions. Beyond the present study, computational modeling can also help us better understand social influence in different modalities, contexts, and social structures.

Background

Particularly informative for us is a study by Salganik, Dodds, and Watts (2006), which focused on the dynamics of social influence in the auditory modality via an artificial song market. Participants were presented with a selection of multiple songs and given the opportunity to download songs of their choosing. When participants had access to the total number of times a song was downloaded by others, Salganik et al. (2006) observed increased inequality in download count between songs. In other words, social context led to greater polarization between the most and least popular songs. This suggests that when people make judgments in social contexts, they tend to amplify information towards extreme ends–a significant consequence of social influence at the population

In L. K. Samuelson, S. L. Frank, M. Toneva, A. Mackey, & E. Hazeltine (Eds.), *Proceedings of the 46th Annual Conference of the Cognitive Science Society.*

level. However, the design of this study was such that there was no system for incorporating a negative evaluation into social information; participants could either contribute positively to the social signal by downloading, or not contribute at all by declining to download. To address this, our graded rating scale allows participants to both positively and negatively affect social information. Salganik et al. (2006) also observed that inequality was enhanced when songs were sorted by popularity, supporting the idea that relative information can influence social dynamics. Our study aims to replicate these findings in the visual modality, using ratings of artworks. We also include additional interface manipulations to address the fact that presenting many stimuli together emphasizes relative differences; in addition to the order of stimuli, we manipulate whether stimuli are presented in groups or individually.

Finally, to model how social information and individual judgments are combined to produce ratings, we use a model for describing multi-sensory signal combination (Ernst & Banks, 2002; Alais & Burr, 2019). This is a weighted linear model that determines the weight to be placed on each signal based on their relative variance, such that each weight is inversely proportional to signal noisiness. By comparing this model against empirical data, we aim to see to what extent it can describe signal combination in social contexts.

Approach

Evaluation Experiments

We conducted experiments via PsyNet (`https://www.psynet.dev`; Harrison et al., 2020), a platform for building complex online studies. PsyNet relies on dallinger (`https://github.com/Dallinger/Dallinger`), an open source platform for deploying large scale experiments. In the *social* condition, pariticipants' average responses to stimuli were passed along as inputs for subsequent participants in a transmission chain structure (Figure 1). This was complemented by an *asocial* control condition where information from previous trials was not propagated. Each experiment involved 10 chains with 10 stimuli assigned to each (i.e. 100 stimuli in total). Each participant entered each chain once, and the order of chains was randomized for each participant. Each chain consisted of 20 iterations (i.e. participants). The number of stimuli and the length of each chain were determined by a pilot study.

Model

We identify the socially-derived component of ratings by fitting our empirical data to a computational model; this not only gives us a concrete value to describe the proportional contribution of social information to an individual judgment, but also helps us compare the strength of influence for different interfaces. We then compare this empirically-derived contribution to an ideal solution for noise-reducing signal combination. This gives us a model-based understanding of the qualitative differences in the process of combining multiple perceptual signals and the process of incorporating social context into a personal judgment.

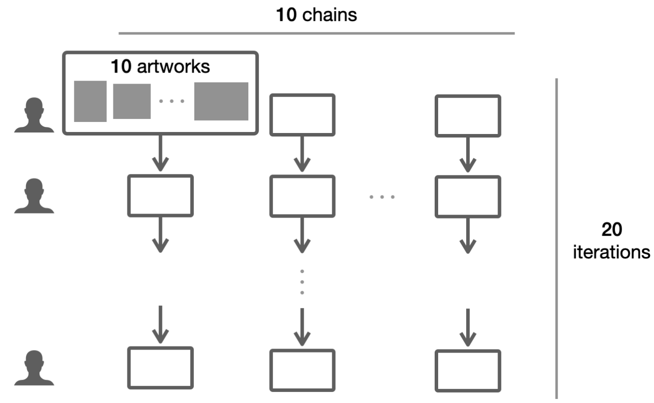

Figure 1: Transmission chain paradigm: participant responses are used as inputs for subsequent participants. There were 10 chains; each chain was 20 iterations long.

We consider a basic model that computes a weighted linear combination of signals. An evaluation X_{rating} is derived from the social information X_o and individual evaluation X_i, where w is the weight placed on social information:

$$X_{rating} = wX_o + (1-w)X_i + n_i \tag{1}$$

We assume the error n_i is an unbiased error from a normal distribution $\mathcal{N}(0, \sigma_n)$ and includes both decision noise and individual differences; see Discussion for an alternative way to integrate individual differences into the model.

In an ideally noise-reducing model, the weight w assigned to the social signal should be inversely proportional to its noisiness; intuitively, the noisier a signal is, the less we should rely on it. This ideal weight can be formulated as:

$$w = \frac{\sigma_i^2}{\sigma_o^2 + \sigma_i^2} \tag{2}$$

where σ_o^2 is the variance of the social signal X_o and σ_i^2 is the variance of the individual signal X_i. This formulation of the weight can largely explain cognitive computations for combining multisensory cues (Ernst & Banks, 2002; Alais & Burr, 2019); it therefore provides a baseline for what signal combination looks like in non-social contexts.

In each experiment, we use the mean rating of previous participants' responses to the same stimulus as social information X_o and the variance of these ratings as σ_o. The participants' individual evaluation X_i and the variance of this signal σ_i^2 are approximated using the ratings of each stimulus from the *asocial* condition. We do not explicitly consider individual differences in the effects of social information, focusing on trends we observe at the group level.

We compare the ideal w from Equation 2 with an empirical w that is derived by solving the regression problem of Equation 1 numerically, with the constraint of keeping w within

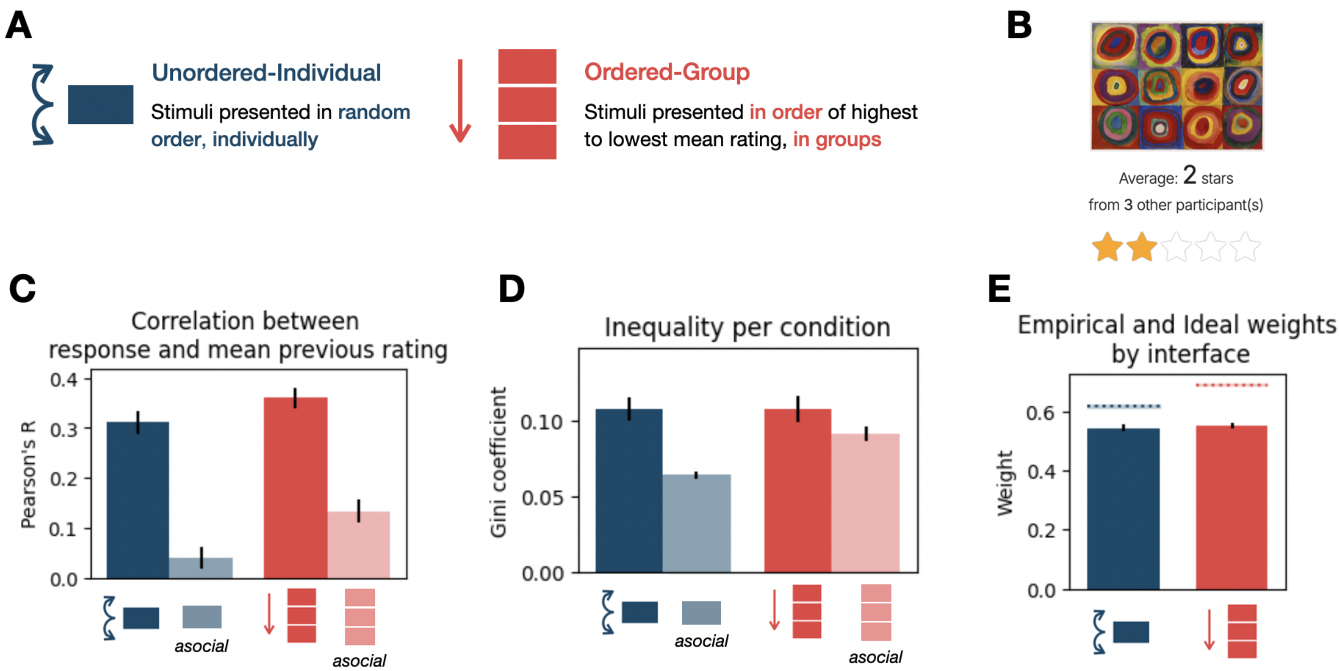

Figure 2: Experiment 1 design and results. (A) Two interfaces were tested to examine how presentation style shapes social influence: the *Unordered-Individual* interface (Experiment 1a) and the *Ordered-Group* interface (Experiment 1b). (B) Example of *Unordered-Individual* interface. (C) In both interfaces, the correlation between current and previous ratings is significantly higher in the presence of social information. (D) Inequality between stimuli by interface and condition. Only the *Unordered-Individual* interface showed significant difference between conditions. (E) Social weights produced by an ideal model are larger than empirically derived weights. For all bar plots, error bars represent one standard error; darker bars represent the social condition and lighter bars represent the asocial.

[0, 1]. By comparing these weights, we aim to understand the strength of social influence and to what extent it can be explained as basic signal combination. Furthermore, this strategy allows us to pinpoint the differences in interfaces in terms of the contribution of the social signal specifically, separate from individuals' pre-existing judgments.

Experiment 1: The Effect of Interface

Methods

Participants We recruited 87 total participants on Prolific, an online recruitment platform (45 participants in Experiment 1a and 42 participants in Experiment 1b). Our criteria for recruitment were that each participant was over 18 years of age, spoke English, and resided in the United Kingdom. All gave informed consent prior to their participation (approved protocol of the Max Planck Ethics Council #2021_42) and passed a color-blindness test (Clark, 1924).

Stimuli Participants evaluated a set of 100 artworks from the WikiArt Emotions Dataset (Mohammad & Kiritchenko, 2018). The works were randomly chosen from a filtered list of artworks that were a) abstract in style, b) paintings, and c) did not include face or body depiction. We limited stimuli to one art style to reduce the effect of preconceived attitudes about different styles. Abstract art in particular was chosen to avoid biases related to semantic content as best as possible.

Procedure After task instructions, participants were shown three random examples of the stimuli. This was followed by the main task of evaluating artworks on their visual appeal using a five-star rating scale (1 star: very unappealing, 5 stars: very appealing). In the *social* condition, participants were shown social information in the form of the mean of previous ratings for each stimulus, rounded to the nearest integer; this was indicated through text and visualization with stars (Figure 2B). In the *asocial* condition, participants were simply presented with an artwork and asked for a rating. Each session included 10 repeat trials consisting of a randomly chosen stimuli set, administered at the end of all trials.

Figure 2A illustrates the two interfaces used in the experiment. We operationalize the absolute and relative nature of these interfaces by manipulating two features in a two-by-two factorial design: Order (stimuli are either ordered by rating to show relative placing, or randomly ordered) and Grouping (stimuli are either presented in groups or individually). In Experiment 1a, stimuli were presented individually, in random order (*Unordered-Individual*). In Experiment 1b, stimuli were presented in sets of 10, and in the social condition, were ordered from highest to lowest mean previous rating (*Ordered-Group*). Experiment 1b most closely resembles the interface used by Salganik et al. (2006), while Experiment 1a is a common interface used for rating experiments in psychology (Likert, 1932; Palmer, Schloss, & Sammartino, 2013). Experiments 1a and 1b were run separately.

Results

We first calculated within-rater consistency by finding the correlation of ratings between original and repeat trials for each participant. The average correlation was $r = .69$ (CI = [.59, .80][1] in Experiment 1a and $r = .62$ (CI = [.54, .71]) in Experiment 1b, supporting the reliability of our ratings.

Figure 2C shows the correlation between given ratings and the rounded mean of previous ratings for that stimulus. In the *social* condition, this value is equivalent to the social information that was presented to the participant. To calculate the *asocial* correlation, we also round the mean to the nearest integer for consistency. In Experiment 1a (*Unordered-Individual*), the correlation for the *social* condition ($r = .31$, CI = [.27, .35]) was significantly higher than that for the *asocial* condition ($r = .04$, CI = [-.004, .08]); $p < .001$, $d = 12.50$. Similarly, in Experiment 1b (*Ordered-Group*), the *social* correlation ($r = .36$, CI = [.32, .40]) was significantly higher than the *asocial* ($r = .13$, CI = [.09, .18]); $p < .001$, $d = 10.30$. This is consistent with what we would expect if social information indeed influences evaluations. The weakly positive correlation in the *asocial* condition is something we would expect given that the correlation between current and previous ratings captures patterns in ratings for each stimuli. The *Ordered-Group* interface produces a significantly higher asocial correlation than the *Unordered-Individual* interface ($p < .05$, $d = 4.14$). As the stimuli were randomly ordered in the asocial condition for both interfaces, this difference seems to be driven by the number of stimuli presented at once.

To better understand stimuli-based patterns of evaluations, we calculated the correlation between ratings for each stimulus across conditions and interfaces. Correlations between the *social* and *asocial* conditions within each interface ranged from $r = .71$ to $r = .84$ ($p < .001$ for all interfaces). Pairwise correlations between different interfaces matched by condition (e.g., comparing the *social* condition in the *Ordered-Group* interface with the *social* condition in the *Unordered-Individual* interface) were between $r = .68$ and $r = .83$ (all $p < .001$). As each condition and interface has a distinct group of participants, these values suggest a pattern in the visual appeal of stimuli that holds across our experimental manipulations and is shared across participants.

We also measured self-reported levels of social influence by directly asking (at the end of the experiment) how often participants took social information into account. In Experiment 1a, seven of the 21 participants in the social condition answered "Not at all", while the rest answered "Sometimes". When we separately analyze the correlation between rating and social information for this group of people ($r = .37$, CI = [.31, .44]), we find no significant difference from the rest of participants' correlation ($r = .29$, CI = [.24, .34]; $p = .06$, $d = 2.67$). In other words, even participants who self-reported no consideration of social information gave ratings that highly correlated with this information. We find the same effect in Experiment 1b. 10 out of 20 participants responded "Not at all" (9 answered "Sometimes"; 1 answered "Most of the time"); participants who answered "Not at all" produced a correlation of $r = .38$ (CI = [.33, .44]) compared to $r = .36$ (CI = [.30, .41]) for the rest of the sample; $p = .49$, $d = 0.90$. In both interfaces, social influence seems to be present regardless of whether participants are aware of it.

[1] All confidence intervals (CIs) reported for correlations are 95% CIs from bootstrapping pairs of rating and mean previous rating.

Inequality Observed inequality is shown in Figure 2D; we adapted Salganik et al.'s (2006) inequality measure by considering the "market share" m_i of each stimulus, defined as the cumulative number of stars received by that stimulus normalized over the set. We then calculated the Gini coefficient $G = \sum_{i=1}^{N}\sum_{j=1}^{N}|m_i - m_j|/2N\sum_{i=1}^{N}m_i$ for each set, where N is the number of stimuli in the set (here, $N = 10$). $G = 0$ indicates perfect equality, while $G = 1$ indicates complete inequality. In Experiment 1a, the mean Gini coefficient was $G = .11$ (CI = [.06, .16]) for the *social* condition and $G = .06$ (CI = [.05, .08]) for the *asocial* condition; $p < .001$, $d = 2.38$. In Experiment 1b, the mean Gini coefficient was $G = .11$ (CI = [.05, .16]) for the social condition and $G = .09$ (CI = [.06, .12]) for the asocial condition; $p = .15$, $d = 0.73$. We again see a difference between interfaces, with only the *Unordered-Individual* interface weakly replicating the inequality dynamic seen by Salganik et al. (2006).

Computational Model Figure 2E shows the difference between the mean weight across all trials for both the empirical and ideal weights. For Experiment 1a, we see that the ideal weight $w = .62$ (CI = [.61, .63], $R^2 = -.03$) is higher than the empirical weight $w = .54$ (CI = [.52, .56], $R^2 = .35$). Similarly, the ideal weight for Experiment 1b is $w = .69$ (CI = [.68, .70], $R^2 = .05$), while the empirical weight is $w = .55$ (CI = [.53, .57], $R^2 = .33$). This discrepancy indicates that people rely less on social information than they should, according to an ideally noise-reducing model.

Experiment 2: Additional Interfaces

Methods

The differences we observed between interfaces in Experiment 1 call for further testing of the effect of each interface feature. We did so by interpolating between the interfaces of Experiments 1a and 1b to produce two new interfaces. Experiment 2a presents 10 stimuli together, in random order (*Unordered-Group*); Experiment 2b presents stimuli individually, in order of rating within each set of 10 (*Ordered-Individual*). A total of 44 participants took part in Experiment 2a and 43 participants in Experiment 2b. As with Experiment 1, all participants gave informed consent and passed a colorblindness test. The stimuli and task were identical to Experiment 1, and Experiment 2a and 2b were run simultaneously.

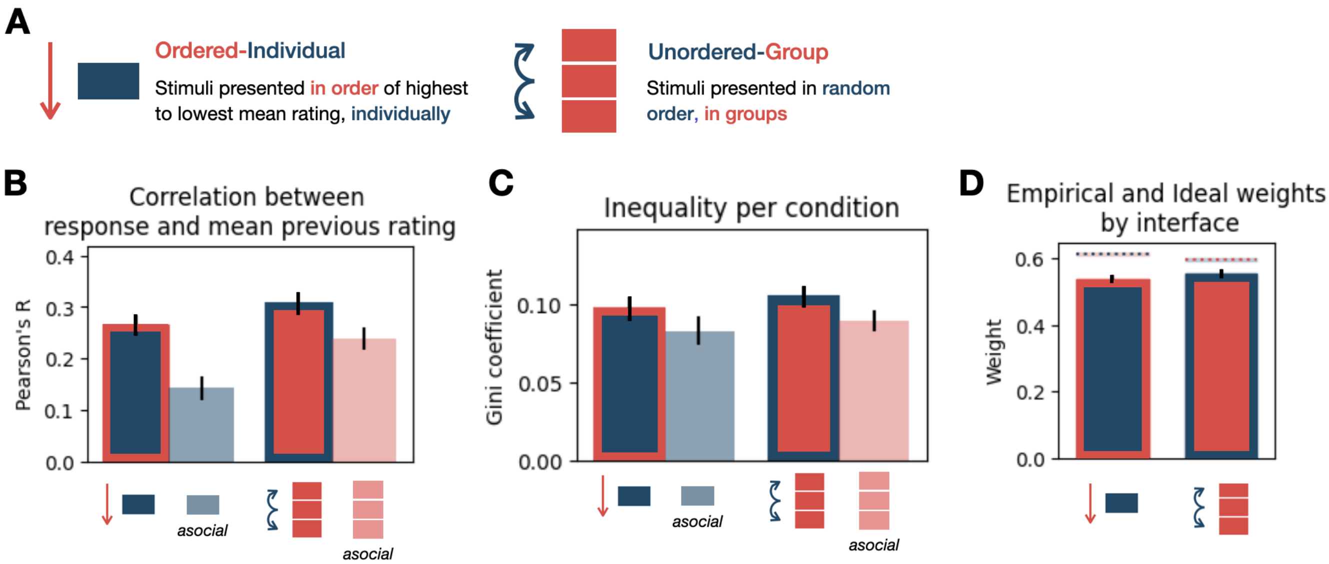

Figure 3: Experiment 2 schematics and results. (A) We interpolate interface features from Experiment 1 to produce an *Unordered-Group* interface (Experiment 2a) and an *Ordered-Individual* interface (Experiment 2b). (B) Social influence by interface; the correlation between current ratings and the mean of previous ratings is higher in the *social* condition for both interfaces than in the *asocial* condition. (C) Inequality between stimuli for each condition, by interface. Neither showed signifcant social effects. (D) The ideal model predicts higher weight on the social signal than can be derived from empirical data. Error bars represent one standard error.

Results

The mean within-rater correlation of ratings was $r = .72$ (CI = [.67, .78]) across both interfaces. This high correlation indicates the reliability of the ratings.

Figure 3C shows that in the *Unordered-Group* interface, the correlation between ratings and the mean of previous ratings in the *social* condition $r = .31$, (CI = [.26, .35]) was significantly higher than in the *asocial* condition $r = .24$, (CI = [.19, .28]); $p = .02$, $d = 3.10$. Similarly, in Experiment 2b (*Ordered-Individual*), the correlation was significantly higher in the *social* condition $r = .26$, (CI = [.22, .31]) than in the *asocial* condition $r = .14$, (CI = [.10, .19]); $p < .001$, $d = 5.53$. This shows that there was significant social influence on evaluations across various interfaces. In addition, the textitasocial correlation in the textitUnordered-Group interface is significantly higher than in the *Ordered-Individual* interface ($p = .002$, $d = 4.33$). This replicates Experiment 1; interfaces with multiple stimuli produce higher asocial correlations than interfaces with individual stimuli.

We also again observe that a self-reported lack of social influence does not correspond to different behavior. In Experiment 2a, eight of the 21 participants in the *social* condition answered "Not at all" regarding how often they took social information into account (12 participants responded "Sometimes"; one responded "Most of the time"). The correlation between rating and social information for those who responded "Not at all" was $r = .27$ (CI = [.21, .34]) compared to $r = .34$ (CI = [.29, .39]) for the rest of participants; $p = .11$, $d = 2.30$. In Experiment 2b, 14 out of 20 participants in the *social* condition answered "Not at all", with five participants answering "Sometimes" and one answering "Most of the time". The correlation between rating and social information for those who responded "Not at all" was $r = .27$ (CI = [.22, .32]) and the correlation for the rest of participants was $r = .25$ (CI = [.16, .33]); $p = .66$, $d = 0.68$. Across all of the interfaces that we tested, we consistently see that social influence affects evaluations regardless of self-reported influence.

Inequality Figure 3C shows the inequality of stimulus ratings in each condition and interface. For the *Unordered-Group* interface, the mean Gini coefficient (i.e. inequality between stimuli) was $G = .11$ (CI = [.06, .15]) in the *social* condition and $G = .09$ (CI = [.05, .13]) in the *asocial* condition; $p = .14$, $d = 0.734$. For the *Ordered-Individual* interface, the mean Gini coefficient of the social condition $G = .10$ (CI = [.08, .11]) and the asocial condition $G = .08$ (CI = [.07, .10]) were also not significantly different ($p = .28$, $d = 0.53$). Thus, neither interface led to significant polarization.

Computational Model We use the same process for deriving empirical and ideal weights as with Experiment 1. Figure 3D shows that the ideal weights for the *Unordered-Group* interface ($w = .60$; CI=[.59, .60], $R^2 = -.09$) and the *Ordered-Individual* interface ($w = .62$; CI=[.61, .62], $R^2 = -.09$) were much higher than the empirical weights for each interface; $w = .56$ (CI=[.53, .58], $R^2 = .31$) for *Unordered-Group* and $w = .54$ (CI=[.52, .56], $R^2 = .28$) for *Ordered-Individual*. This is consistent with our findings from Experiment 1; human participants rely less on social information than would be predicted from an ideal noise-reducing model.

Discussion

We examined two important aspects of understanding social influence in aesthetic evaluations. First, we explored effects of the interface by which information is presented. Second, we explored the mechanism behind social influence by comparing an ideal model to empirical data.

Across all of the interfaces tested, we observed that given knowledge of others' ratings, people produce evaluations that are relatively similar to those ratings. We also observed small differences in effect size between interfaces. However, these differences do not seem to originate from differences between relative and absolute social information, as they are also present in asocial contexts. Namely, interfaces that presented groups of stimuli at once produced highly correlated asocial ratings between participants, compared to interfaces with individual stimuli. This suggests that the effect is largely driven by the availability of comparisons across groups of stimuli, rather than by social information. This is important because, in real-life applications such as recommendation systems, dynamics often associated with social feedback may actually arise from underlying stimulus relationships that are emphasized in grouped interfaces.

The interfaces also showed differing inequality effects, with only one of four interfaces replicating the effect observed by Salganik et al. (2006). It would be difficult to draw strong conclusions about the inequality effect from this, however, because the interface used by Salganik et al. (2006) included significant differences from ours—e.g., participants in their study did not have to evaluate every song in the list, facilitating polarization between popular and unpopular songs. In addition, our study involved far fewer iterations and stimuli, which may have under-powered our results. We can, however, take this result as further suggesting that differences in interface can paint different pictures about the population-level effects of social information. This emphasizes the value of systematically testing out multiple interfaces for the same experimental task.

When we examine social influence via a computational model, we found that the ideally noise-reducing weight on social information is higher than the empirically derived weight, across all interfaces. We can infer from this that humans underestimate the reliability of the social signal; in other words, people believe that social information has limited relevance for their own aesthetic evaluations. Similar discounting of others' opinions can be seen in cases of advice-taking, possibly explained by the fact that people have access to their own reasoning for an evaluation, but not to others' (Yaniv & Kleinberger, 2000). This suggests that while studies on objective judgments have often shown strong social influence, subjective judgment-making may involve qualitatively different processes for taking social information into account. We could also think of individual evaluations as a prior that is updated by observing social information. Under this framing, participants may have relied heavily on the prior because they did not have access to the true distribution of ratings.

Limitations and Future Directions

In our computational approach, we showed that an ideal model is inadequate for describing the way people take social information into account in aesthetic evaluations. What kind of model, then, can better approximate this process? To address this, one strategy would be to construct a model that takes into account only the information that is available to participants. For example, we can describe social information X_o from the participants perspective as a combination of their individual evaluation X_i and the extent to which others are different from them d_i. This difference d_i would then update throughout the course of the experiment as participants gather more observations about others' evaluations. This changes how we formalize the noisiness in social information; rather than taking the veridical variance, as we did in our ideal model, this would use the difference in self and others' tastes as an estimate of perceived reliability.

Our experimental paradigm can also be applied to related research questions. For example, future work might focus on how individual differences manifest in social influence dynamics; susceptibility to social influence may very well differ among individuals (Cascio, Scholz, & Falk, 2015; Oyibo & Vassileva, 2019) potentially creating dynamics that we may have missed from our population-level perspective. Another exciting direction would be to introduce artificial social information to better understand the causal relationship between social information, social structure, and subjective judgment.

Conclusion

Taken together, our results provide an overview of how different interfaces may affect the strength of social influence in aesthetic evaluations. Furthermore, our model of integrating social information and individual judgment shows us that the computational mechanisms behind this process cannot be explained as a simple noise-reducing combination of signals. This deepens our understanding of social learning and cultural evolution on a greater scale. More broadly, our experiments showcase the utility of using large-scale online experiments with adaptable and easily reproducible code. The resulting data in turn facilitate testing of computational models.

Our results show how we can study social influence on subjective judgments in a way that links the computational processes at the individual level with dynamics at the population level. This is not limited to our setting of visual aesthetics; we can adapt this paradigm for studying other aesthetic modalities such as music, or for subjective judgments such as morality or political opinion. Understanding social influence in these domains will deepen our understanding of social cognition, judgment, and decision-making as a whole.

Acknowledgments We would like to thank Peter Harrison, Pol van Rijn, and Frank Höger for their help and support.

References

Alais, D., & Burr, D. (2019). Cue Combination Within a Bayesian Framework. In A. K. C. Lee, M. T. Wallace, A. B. Coffin, A. N. Popper, & R. R. Fay (Eds.), *Multisensory Processes* (Vol. 68, pp. 9–31). Springer International Publishing. (Series Title: Springer Handbook of Auditory Research)

Asch, S. E. (1951). Effects of group pressure upon the modification and distortion of judgments. In *Groups, leadership and men; research in human relations* (pp. 177–190). Oxford, England: Carnegie Press.

Cascio, C. N., Scholz, C., & Falk, E. B. (2015, June). Social influence and the brain: persuasion, susceptibility to influence and retransmission. *Current Opinion in Behavioral Sciences*, *3*, 51–57.

Clark, J. H. (1924). The Ishihara Test for Color Blindness. *American Journal of Physiological Optics*, *5*, 269–276.

Ernst, M. O., & Banks, M. S. (2002, January). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature*, *415*(6870), 429–433. (Number: 6870 Publisher: Nature Publishing Group)

Harrison, P., Marjieh, R., Adolfi, F., van Rijn, P., Anglada-Tort, M., Tchernichovski, O., . . . Jacoby, N. (2020). Gibbs Sampling with People. In *Advances in Neural Information Processing Systems* (Vol. 33, pp. 10659–10671). Curran Associates, Inc.

Henrich, J., & Gil-White, F. J. (2001, May). The evolution of prestige: freely conferred deference as a mechanism for enhancing the benefits of cultural transmission. *Evolution and Human Behavior*, *22*(3), 165–196.

Kendal, R. L., Boogert, N. J., Rendell, L., Laland, K. N., Webster, M., & Jones, P. L. (2018, July). Social Learning Strategies: Bridge-Building between Fields. *Trends in Cognitive Sciences*, *22*(7), 651–665.

Likert, R. (1932). A technique for the measurement of attitudes. *Archives of Psychology*, *22 140*, 55–55.

McNamara, J. M., Fawcett, T. W., & Houston, A. I. (2013, May). An Adaptive Response to Uncertainty Generates Positive and Negative Contrast Effects. *Science*, *340*(6136), 1084–1086. (Publisher: American Association for the Advancement of Science) doi: 10.1126/science.1230599

Mohammad, S., & Kiritchenko, S. (2018, May). WikiArt Emotions: An Annotated Dataset of Emotions Evoked by Art. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC 2018).* European Language Resources Association (ELRA).

Moore, D. A., & Klein, W. M. P. (2008, September). Use of absolute and comparative performance feedback in absolute and comparative judgments and decisions. *Organizational Behavior and Human Decision Processes*, *107*(1), 60–74. doi: 10.1016/j.obhdp.2008.02.005

Oyibo, K., & Vassileva, J. (2019, September). The relationship between personality traits and susceptibility to social influence. *Computers in Human Behavior*, *98*, 174–188.

Palmer, S. E., Schloss, K. B., & Sammartino, J. (2013). Visual Aesthetics and Human Preference. *Annual Review of Psychology*, *64*(1), 77–107.

Salganik, M. J., Dodds, P. S., & Watts, D. J. (2006, February). Experimental Study of Inequality and Unpredictability in an Artificial Cultural Market. *Science*, *311*(5762), 854–856. (Publisher: American Association for the Advancement of Science)

Soll, J. B., & Larrick, R. P. (2009). Strategies for revising judgment: How (and how well) people use others' opinions. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 780–805.

Verpooten, J., & Dewitte, S. (2017, March). The Conundrum of Modern Art. *Human Nature*, *28*(1), 16–38.

Yaniv, I., & Kleinberger, E. (2000, November). Advice Taking in Decision Making: Egocentric Discounting and Reputation Formation. *Organizational Behavior and Human Decision Processes*, *83*(2), 260–281. doi: 10.1006/obhd.2000.2909