Acta Psychologica 208 (2020) 103082

Contents lists available at ScienceDirect

Acta Psychologica

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

On the relation between perceived stability and aesthetic appreciation

Martin G. Fillinger*, Ronald Hübner

University of Konstanz, Germany

ARTICLE INFO

Keywords:
Aesthetics
Preference
Stability
Perceptual balance
Artwork

ABSTRACT

Perceived stability is an important feature of pictures with respect to their aesthetic appreciation. Pictures whose composition is perceived as stable are usually liked more than those with unstable arrangements. However, there are exceptions. In a recent study, we found that unstable Japanese calligraphies were preferred to stable ones. From this result, we hypothesized that instability is liked when it implies movement. Therefore, we systematically tested these two types of instability. In our first experiment, we used multiple-element pictures of varying stability as stimuli and show that perceived instability has a negative effect on liking. In a second experiment, we used dynamic paintings by the artist K.O. Götz, which largely vary in implied movement. As expected, for these dynamic pictures, instability was positively related to liking. Taken together, our findings indicate that perceived instability reduces the aesthetical appreciation of a picture unless it implies movement.

1. Introduction

How well a picture is liked can depend on various factors. Some of them are related to the picture's perceptual features, others to its semantic content. An example of the first type is *perceptual balance*, whose importance has been proposed by art theorists (e.g., Arnheim, 1982; Bouleau, 1980; Kandinsky, 1926/1979; Ross, 1907) and confirmed by empirical researchers (e.g., Locher, 2006; Locher, Stappers, & Overbeeke, 1998; McManus, Edmondson, & Rodger, 1985). The art theorist Ross (1907), for example, considered balance as part of harmony. He wrote that a balanced arrangement of elements "is a Harmony of Positions due to the coincidence of two centers, the center of the attractions and the center of the framing" (p. 24). Whereas the geometric center of a frame can easily be located, it is unclear how the center of attractions can be determined. A widely used approach in this respect is to use mechanical balance as a metaphor. It is assumed that each element in a picture has a certain perceptual weight that depends on its low-level features such as color, size and form, as well as on its semantic content (e.g., Arnheim, 1982). However, the problem remains to calculate the weights. In the simplest case, the weight of an element is determined by its area and gray level. The center of the attraction (perceptual weights) is then computed analogously to the center of mass in mechanics.

The mechanical approach has also been used to compute objective measures of balance. For instance, for the DCM (Deviation of the Center of Mass), a measure proposed by Hübner and Fillinger (2016), it is assumed that the perceptual weight of each pixel is related to its gray level. For the stimuli used in this study, it was assumed that the perceived "mass" of a pixel increases from white to black, i.e., a dark pixel is perceived as heavier than a brighter one. The center of mass in a picture can then easily be computed analogously to mechanics (McManus, Stöver, & Kim, 2011). The DCM score, i.e., the degree of balance is finally defined, as suggested by Ross (1907), by the distance of the center of mass from the geometric center of the picture. A similar measure is the APB (Assessment of Preference for Balance), proposed by Wilson and Chatterjee (2005), who also assumed that the weight of each pixel is related to its gray level. The APB score is defined as the average of eight symmetry measures over the four axes of a picture (horizontal, vertical and the two diagonals).

At least for simple pictures that included only basic and unrelated geometrical elements (e.g., circles, or squares), these measures have successfully been applied to predict balance ratings and liking. Moreover, Thömmes and Hübner (2018) analyzed about 700 architectural photographs and found that for those depicting a real scene (with a 3D appearance), the scores of the objective measures of balance (DCM and APB) significantly predicted the number of Instagram likes. The more balanced a picture was, the more likes it obtained.

However, there are also negative results. Gershoni and Hochstein (2011), for instance, used Japanese calligraphies as stimuli and found that the APB failed to predict balance ratings. Recently, Fillinger and Hübner (2018) replicated this result. Moreover, they showed that for these pictures, balance ratings were also unrelated to liking ratings. This suggests that there are different types of balance. Indeed, further data collection and analyses revealed that the liking of calligraphies

* Corresponding author at: Department of Psychology, University of Konstanz, P.O. Box 29, 78457 Konstanz, Germany.
E-mail address: mgfillinger@gmail.com (M.G. Fillinger).

https://doi.org/10.1016/j.actpsy.2020.103082
Received 4 November 2019; Received in revised form 9 April 2020; Accepted 29 April 2020

was affected by perceived stability, which is considered as different from but closely related to balance (Ross, 1907).

If we go back in history, then, as far as we know, the difference between balance and stability has first been examined systematically by Pierce (1894, 1896), who at that time, was a graduate student in Hugo Münsterberg's lab at Harvard University. In his studies, Pierce wanted to investigate perceptual balance by asking participants to adjust the horizontal position of a movable object (e.g., a line) on one side of a display such that this side was aesthetically equal to the opposite side, where another object (e.g., a line of different size) was fixed at a certain position. At least under some conditions, adjustments were made in accord with mechanical balance. Interestingly, however, Pierce also rotated the displays by 90° and found that in this case, the adjustments changed. For vertical layouts, gravitational stability was more important than balance. Accordingly, pictures were preferred when they had more weight in their lower half rather than in their upper one. Thus, it seems that pictures are preferred whose elements are arranged in a gravitationally stable way. This is in line with more recent research by Friedenberg (2012), who observed that triangular shapes perceived as unstable (e.g., because they stood on one of the edges instead of one of the baselines), were rated as less attractive.

If stability usually affects liking positively (i.e., the more stable its composition the more a picture is liked), then one can ask why the relation was reversed in our previous study with Japanese calligraphies. Or, more generally, why does instability increase liking in some pictures and decrease it in others? One possible explanation is that there are (at least) two types of stability. One type is gravitational stability (van der Helm, 2015), which is preferred because it prevents damage and injuries. Due to corresponding associations, this preference is also generalized to the content of pictures.

However, what is the other type? We hypothesized that it is related to the dynamics and movement implied by the objects in a picture. These features also play a role in related areas. In a study with human figures (Friedenberg, Keating, & Liby, 2012), for instance, it has been shown that estimates of the center of mass are biased towards the direction of the implied movement. Moreover, in Chinese art theory, it is assumed that a brushstroke expresses the painter's emotion, which also holds for calligraphies (Dubal et al., 2014). Thus, expressive brushstrokes represent dynamics and imply movement, while at the same time they may appear unstable. Nonetheless, due to the implied movement (Mather & Sharman, 2015; Osaka, Matsuyoshi, Ikeda, & Osaka, 2010), which usually evokes positive emotions, the aesthetic appeal of corresponding pictures is high. That emotions play an essential role in the processing of aesthetic stimuli has widely been assumed (Menninghaus et al., 2019).

This conjecture is also supported by another of our recent studies (Hübner & Fillinger, 2019), where we used pictures from the Visual Aesthetic Sensitivity Test (VAST; Götz, 1985) as stimuli. The VAST pictures can be categorized into single-element, multiple-element and dynamic pictures. We found that, overall, there was a negative correlation between stability and liking. In particular, the dynamic pictures were rated as highly unstable but were nevertheless liked most. The multiple-element pictures were liked less and were rated, on average, as more stable. Interestingly, within this category, there was a positive relation between stability and liking.

Taken together, the considered results suggest that at least two types of instability can be differentiated: one type is associated with gravitation and is disliked, whereas the other type is associated with movement and is liked, presumably, because of the involvement of emotion (Dubal et al., 2014).

The aim of the present study was to systematically investigate these two types of instability. In our first experiment, we applied multiple-element pictures. For these stimuli, we hypothesized that instability has a negative effect on aesthetic appreciation, as suggested by results obtained with similar pictures in one of our previous studies (Hübner & Fillinger, 2019). In the second experiment, we presented dynamic stimuli, which were similar to untypical Japanese calligraphy (Fillinger & Hübner, 2018). We expected that for these dynamic pictures instability is liked, because they imply movement.

2. Experiment 1

In our first experiment, we wanted to show that perceived gravitational instability can have a negative effect on aesthetic appreciation. For this objective, we constructed a basic set of four pictures, each showing three rectangles and three colored decorative elements (see Table 1). Although these basic stimuli already differed in stability, this property was further varied by rotating the stimuli. Participants had to rate balance and stability for each of these stimuli, and they were also asked to rate how much they liked the pictures. For comparison with Experiment 2, we later also collected movement and emotionality ratings.[1]

Furthermore, we computed APB and DCM scores for the pictures. Because symmetry might also be important for explaining the liking of our stimuli (e.g., Jacobsen & Höfel, 2002; Tinio & Leder, 2009), we additionally computed vertical mirror symmetry scores, henceforth referred to as *mirror symmetry* (MS), with the method described in Hübner and Fillinger (2016).

We expected similar results as in Hübner and Fillinger (2019) for multiple-element pictures. Specifically, balance and stability should be positively correlated. Most importantly, however, participants should prefer balanced and stable pictures.

2.1. Method

2.1.1. Participants

Eighty-seven persons (26 males, mean age 22.4, $SD = 5.67$) were recruited via an online system (ORSEE; Greiner, 2015) for participation in the online experiment. All participants received a 3 € voucher as an incentive. The study was carried out in accordance with the ethical guidelines of the University of Konstanz and the Declaration of Helsinki. Participants were informed of their right to abstain from participation in the study or withdraw consent to participate at any time without reprisal. The experiment could only be started after the informed consent was accepted by marking a checkbox on the first page of the online experiment.

2.1.2. Stimuli

Stimuli were 32 pictures (500 × 500 pixels) composed of three black rectangles, whose formations were more or less stable. The pictures also contained three small decorative colored elements (for examples see Table 1), which were inspired by the artwork of Kasimir Malewitsch (1878–1935) and should make the stimuli look like abstract artwork. The stimulus set was based on four basic stimuli (#1, 9, 17, and 25 in Table 1) that were thought to differ in stability. Additional stimuli were created by rotating (30, 90, 150, 180, 210, 270, and 330°) these stimuli (all stimuli are accessible on https://osf.io/ebdah/). The stimuli were presented at the center of the display on a gray background. Stimulus presentation and response registration were controlled by SoSci Survey (Leiner, 2019).

For the stimuli we also computed objective balance (APB and DCM) and mirror symmetry (MS) scores. The APB ranged from 43.2 to 51.1 ($M = 47.3$, $SD = 2.18$), the DCM from 3.77 to 12.8 ($M = 7.89$, $SD = 3.08$), and MS from 6.20 to 17.1 ($M = 11.0$, $SD = 2.45$).

2.1.3. Procedure

The online experiment started with an instruction that informed the participants about the task. Subsequently, two blocks of trials were administered. In the first block, all participants rated how much they

[1] We thank one of the reviewers for this helpful suggestion.

Table 1
Results of Experiment 1.

#	**1**	**5**	**7**	**27**	**3**	**8**	**29**	**31**
S	93	83	71	68	66	63	63	58
L	55	47	49	50	47	52	47	49
B	83	78	68	61	64	59	58	55
M	18	24	26	30	25	38	30	33
E	35	32	34	28	34	34	26	29

#	**17**	**9**	**2**	**15**	**11**	**25**	**19**	**23**
S	57	55	52	52	51	50	50	50
L	40	44	50	41	41	46	43	42
B	43	43	58	45	50	58	43	40
M	43	45	47	44	41	34	48	45
E	23	29	35	26	30	27	25	27

#	**4**	**6**	**13**	**32**	**21**	**30**	**28**	**26**
S	47	46	45	40	40	38	34	34
L	45	45	41	48	43	45	44	48
B	54	55	40	46	37	50	47	48
M	45	50	46	51	47	54	56	54
E	33	35	24	28	25	28	25	27

#	**24**	**16**	**18**	**10**	**22**	**14**	**12**	**20**
S	28	27	26	26	22	22	21	20
L	41	38	37	37	40	40	39	36
B	32	34	27	31	28	34	32	27
M	71	67	69	66	68	66	70	67
E	28	26	28	29	25	25	26	24

Note. The thumbnails of the 32 stimuli, shown in the first row, are ordered by stability. Row two (#) shows the identification number of the pictures. The corresponding mean ratings for stability (S), liking (L), balance (B), movement (M), and emotionality (E) are listed in the rows below.

liked the stimuli (from “I do not like it” to “I like it”). The second block differed between the participants. About half of the participants (44) rated how well the stimuli were balanced (from “not balanced” to “balanced”), whereas the remaining participants (43) rated how stable the pictorial elements were arranged (from “unstable” to “stable”). The ratings were entered by clicking on a visual analogue scale, which ranged from 0 to 100 (not visible for participants). Shortly after each response, the next stimulus was displayed. Altogether, the experiment comprised 64 trials (2 × 32 trials) and lasted about 7 min.

2.1.4. *Additional ratings*

For comparison between Experiment 1 and 2, a further rating experiment was conducted later to collect movement and emotionality ratings for the stimuli of Experiment 1, and balance ratings for the stimuli of Experiment 2. Ninety-one (14 males, mean age 24.6, *SD* = 4.90) persons participated under the same conditions as in Experiment 1 and 2 and received a 3 € voucher as an incentive. The participants were randomly assigned to two groups of 47 and 44 persons, respectively. The first group started with emotionality ratings in the first block, followed by movement ratings in a second block (for further details see Experiment 2), and balance ratings in the third block. In contrast, the second group started with balance ratings followed by emotionality and movement ratings. In each block, the stimuli were presented in random order. Overall, the three blocks lasted about 12 min.

While the additional emotionality and movement ratings were analyzed in the present experiment, the balance ratings were considered in the second experiment.

2.2. *Results*

Mean ratings for liking, balance, stability, movement, and emotionality were 44.0 (*SD* = 4.75), 47.8 (*SD* = 14.3), 46.9 (*SD* = 18.3), 47.6 (*SD* = 15.0), and 28.5 (*SD* = 3.68), respectively. The mean ratings for each picture are presented in Table 1. Correlational analyses revealed significant correlations between all rated properties (see Table 2). However, correlations between movement and the other ratings were negative. As expected, and most important, there was a positive relation between stability and liking, which is also shown in Fig. 1. Concerning the objective measures of balance, there were merely significant correlations of balance and liking with the DCM scores. Mirror symmetry was positively correlated with both balance and stability ratings, and negatively with movement ratings.

Table 2
Correlations between mean ratings for liking, balance, stability, movement, and emotionality in Experiment 1. The correlations between the mean ratings and objective measures of balance (APB and DCM) and mirror symmetry (MS) are also shown.

	Balance	Stability	Movement	Emotionality	APB	DCM	MS
Liking	.866***	.763***	−.727***	.641***	−0.333	−.429*	0.319
Balance	–	.912***	−.890***	.687***	−0.264	−.422*	.460**
Stability	–	–	−.946***	.566**	−0.104	−0.195	.443*
Movement	–	–	–	−.453**	0.217	0.311	−.513**
Emotionality	–	–	–	–	0.036	−0.156	0.318

Note. APB = Assessment of Preference for Balance; DCM = Deviation of the Center of Mass; MS = Mirror symmetry.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

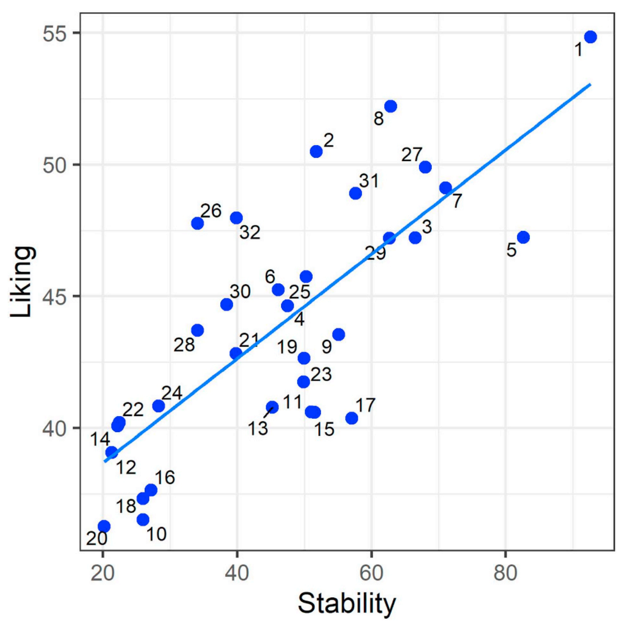


Fig. 1. Relation between stability and liking ratings in Experiment 1. The line represents the corresponding linear regression line. Each data point represents the picture according to its number (see Table 1).

2.3. Discussion

In this experiment, we used multi-element stimuli of varying stability to investigate the relationship between balance, stability, movement, emotionality, and liking. As expected, stability was positively related with liking ratings (see Fig. 1). The same held for balance and liking. Thus, our results support the notion that pictures showing balanced and gravitationally stable compositions are preferred (Friedenberg, 2012; Pierce, 1896). Accordingly, there was also a strong relationship between balance and stability, which replicates results from our previous studies (Fillinger & Hübner, 2018; Hübner & Fillinger, 2019). The fact that this latter correlation was also quite high, indicates that the concepts of balance and stability can be rather similar. However, it is also clear from our results that they differ: Whereas the DCM scores significantly correlated with balance ratings, the correlation with stability ratings was considerably lower and not significant.[2] Moreover, balance ratings correlated much lower with the DCM scores than with stability. This could mean that the balance ratings reflect a mixture of both the usual mechanical balance as well as gravitational stability.

Interestingly, the DCM scores also significantly correlated with liking, which means that a picture was liked more the closer the center of mass was located to the picture's geometric center. The fact that the DCM correlated with both balance and liking is different from our previous results with Japanese calligraphies (Fillinger & Hübner, 2018). The APB scores did not correlate with any rating.

Furthermore, stability, balance, and liking showed a negative relation with movement and a positive one with emotionality, which supports the notion of gravitational stability (van der Helm, 2015). The more pictorial elements imply movement, the more instable looks the picture. Because instability is associated with damage and injuries, this could explain the negative relation between movement and emotionality ratings.

Mirror symmetry showed no significant correlation with liking, which is at odd with previous research (e.g., Jacobsen & Höfel, 2002; Tinio & Leder, 2009). A reason could be that stability and balance dominated the liking of our stimuli. Because stable and balanced images are more symmetrical than moved ones, mirror symmetry correlated positively with stability and balance, and negatively with movement.

Taken together, the results support our hypothesis that, in multiple-element pictures, gravitational stability is strongly related to perceived balance, but also differs from it. Moreover, both features are positively related to liking.

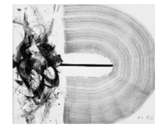
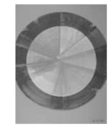
3. Experiment 2

In the previous experiment with multi-element pictures, instability was not liked. These results are in line with art theory (Liu, Dong, Zhang, & Jiang, 2017), as well as with early (Pierce, 1896) and recent (Friedenberg, 2012; Hübner & Fillinger, 2019) experiments. However, the result is different from that observed for Japanese calligraphies. For these stimuli, instability was preferred (Fillinger & Hübner, 2018; Hübner & Fillinger, 2019). In the present experiment, we not only intended to replicate this latter result with a different type of stimuli, but also to examine further the reasons for the inconclusive outcomes. For this objective, we used artwork by Karl Otto Götz (1914–2017), who was one of the most important members of the German Art Informel movement. Götz is known for his explosive and complex abstract forms consisting of a mixture of organic and geometric elements.

A selection of Götz's paintings was rated for perceived stability and liking, but also for movement and emotionality. For comparison with the results of our first experiment, we later also collected balance ratings (see Section Additional Ratings in Experiment 1). We expected that for these stimuli, stability is negatively related to liking and that this relation depends on movement and emotionality. If instability is related to implied movement, then it is presumably liked, because movement is associated with positive emotions (Dubal et al., 2014). Irrespective of their valence, aesthetic emotions and their perceived intensity are understood as one source of liking (Menninghaus et al., 2019). Accordingly, in Experiment 2, we asked participants how strong their emotions are.

[2] It should be noted that the smaller the DCM score, the closer the center of mass is to the center of the frame.

Table 3
Example stimuli used in Experiment 2 with diametrically opposite movement and stability ratings.

					
#	**85**	**9**	**91**	**96**	**16**
M	81	73	58	43	24
S	30	33	35	59	87
L	55	54	42	36	38
E	55	46	40	33	30
B	37	49	57	47	84

Note. The bold numbers in row two (#) represent the corresponding picture numbers. The following rows show movement (M), stability (S), liking (L), emotionality (E), and balance (B) ratings, respectively.

3.1. Method

Before describing the method of the main experiment, we report two preliminary studies. The first one served for selecting an appropriate set of pictures as stimuli. The second one was conducted to rule out a possible confound of movement with curvature.

3.1.1. Preliminary study 1

For this preliminary study, we selected 100 paintings by Karl Otto Götz, which he created between 1936 and 2010. Forty-two persons (9 males; mean age 22.9, $SD = 2.70$) were recruited for the study that was performed under the same ethical standards as the previous experiment. They received a 3 € voucher for their participation. The procedure was similar to the previous experiment. After a short instruction, gray-level pictures of the 100 paintings were presented in random order. Two visual analogue scales located below the picture were used to enter stability (from “unstable” to “stable”) and movement (from “static” to “dynamic”) ratings. There was no time limit. Overall, the study lasted about 10 min.

The mean stability and movement ratings were 45.3 ($SD = 7.10$) and 41.9 ($SD = 20.8$), respectively. The two ratings were strongly negatively correlated ($r = -0.870$, $p < .001$), as expected. For the main study, we selected a representative set of 44 pictures (for example stimuli see Table 3; the 44 selected paintings are accessible on https://osf.io/ebdah/) that almost equally covered the range of perceived stability as well as that of perceived movement. For the selected set, the correlation between stability and movement ratings was $r = -0.896$, $p < .001$. For the final 44 stimuli we also computed objective balance (APB and DCM) and mirror symmetry (MS) scores. The APB ranged from 4.60 to 26.7 ($M = 12.1$, $SD = 4.30$), the DCM from 1.70 to 32.4 ($M = 10.9$, $SD = 6.50$), and MS from 6.20 to 66.7 ($M = 31.5$, $SD = 14.1$).

3.1.2. Preliminary study 2

Because the elements in the selected stimuli (especially the more dynamic ones) are more curved than the Japanese calligraphies used in Fillinger and Hübner (2018), or the picture elements in Experiment 1, and it is known that curvature affects liking (e.g., Bar & Neta, 2006; Gómez-Puerto, Munar, & Nadal, 2016; Silvia & Barona, 2009), we wanted to control for curvature. That is, we wanted to exclude that pictures implying movement are liked more, simply because their elements are also curved. Therefore, we conducted a study in which fifteen participants (4 males; mean age 24.3, $SD = 4.80$) had to rate curvature versus angularity for the selected set of stimuli on a visual analogue scale (from “curved” to “angular”). The recruitment and ethical standards were similar to Preliminary Study 1. Moreover, stimuli were presented under the same conditions as in the previous experiment. The study lasted about seven minutes and the participation was compensated in exchange for course credit. As result, curvature did neither correlate with the stability ratings, $r = -0.109$, $p = .480$, nor with the movement ratings, $r = -0.137$, $p = .375$, from the first preliminary study.

3.1.3. Participants

Ninety-two persons (18 males; mean age 24.7, $SD = 6.30$) were recruited in the same way as in the previous experiments for participation in the online experiment. None of them had participated in the preliminary studies. All other participation criteria were the same as in Experiment 1.

3.1.4. Procedure

The participants were randomly assigned to four groups of 23 persons each. One group started with liking ratings in the first block, followed by movement ratings (from “static” to “dynamic”) in a second block. Another group started with liking ratings, followed by stability ratings. A third group started with emotionality ratings, followed by movement ratings and a final group started with emotionality ratings, followed by stability ratings. For the emotionality ratings, we asked participants to indicate how strongly a picture provokes emotions (from “not at all” to “very strongly”). Similar to Experiment 1, the ratings were entered by clicking on a visual analogue scale. In each block, the gray-level pictures of the 44 selected paintings were presented in random order. There was no time limit. Overall, the two blocks lasted about 10 min.

3.2. Results

Mean liking, emotionality, stability, balance, and movement ratings were 46.2 ($SD = 7.70$), 44.2 ($SD = 9.20$), 44.9 ($SD = 14.7$), 47.8 ($SD = 11.2$), and 56.7 ($SD = 16.2$), respectively. For further analyses, we averaged the ratings across participants for each painting. Example stimuli and their mean ratings can be seen in Table 3. First, it should be noted that the stability as well as the movement ratings were strongly correlated with the respective ratings from the first preliminary study ($r = 0.756$, $p < .001$, for stability, and $r = 0.837$, $p < .001$, for movement).

Further correlational analyses revealed the following significant correlations between the following variables (see Table 4): first, there was a positive relationship between liking and emotionality; second, stability ratings correlated negatively with liking, emotionality, and movement. Movement was positively correlated with liking and with emotionality. Balance was positively correlated with stability, and negatively with movement as well as with emotionality. Importantly, liking was not significantly correlated with both balance and the curvature ratings from the Preliminary Study 2, ($r = 0.198$, $p = .198$). Objective measures for balance (APB and DCM) and mirror symmetry

Table 4
Correlations between the different ratings and objective measures in Experiment 2.

	Balance	Stability	Movement	Emotionality	APB	DCM	MS
Liking	−0.134	−.311*	.549***	.514***	0.296	0.162	0.005
Balance	–	.765***	−.609***	−.502**	−0.070	−0.094	0.008
Stability	–	–	−.843***	−.626***	0.113	0.196	−0.213
Movement	–	–	–	.826***	0.038	−0.195	0.296
Emotionality	–	–	–	–	0.106	−0.084	0.275

Note. APB = Assessment of Preference for Balance; DCM = Deviation of the Center of Mass; MS = Mirror symmetry.
* $p < .05$.
** $p < .01$.
*** $p < .001$.

Table 5
Results of the multiple regression analysis with liking ratings as dependent variable (DV).

DV	Predictor	B	SE	β	*p*	*F*	R^2	RΔ
Liking	–	–	–	–	–	12.7***	0.382	0.285
	(Constant)	46.2	0.940	–	<.001	–	–	–
	Stability	0.277	0.120	0.527	.026	–	–	–
	Movement	0.476	0.109	0.994	<.001	–	–	–

Note.
*** $p < .001$.

(MS) showed no significant correlation with the ratings.

The relation between liking, stability, and movement was further analyzed by multiple linear regression (see Table 5), with liking as the dependent variable, and stability and movement as independent variables (grand mean centered). Due to the strong correlation between stability and movement, collinearity was measured by the variance inflation factor. However, a relatively moderate value ($VIF = 3.46$) indicates that collinearity was not a concern. As a result, stability, as well as movement, significantly accounted for the variance of liking. However, the regression coefficient of stability was positive. This shows that, if the effect of implied movement is controlled for, stability still has an effect on liking, albeit a positive one (see Fig. 2).

3.3. Discussion

In this experiment, we tested the hypothesis that perceived instability, if related to implied movement, increases the liking of an image. For this objective, we used artworks of Karl Otto Götz, which largely varied in implied movement. As expected, pictures that were rated as unstable were liked more than those rated as stable. This is in line with observations in our previous studies (Fillinger & Hübner, 2018; Hübner & Fillinger, 2019). Here, however, the relation was more direct and systematic.

As in Experiment 1, the participants in this experiment rated the emotionality, implied movement, and balance of the pictures. As hypothesized, implied movement was responsible for the negative relation between stability and liking. If the variance accounted for by movement was taken out from the liking ratings, then stability correlated again positively with liking. Similar to a previous study with dynamic pictures (Hübner & Fillinger, 2019), the results showed that balance is closely related to stability. This close relationship is also reflected by similar correlative relations with movement and emotionality. Balance as well as stability correlated negatively with movement and emotionality, respectively.

Furthermore, the objective measures of balance and mirror symmetry showed no significant correlation with the ratings. This indicates that the concepts of these measures did not play any role in predicting our ratings.

Thus, taken together, our results suggest that there are two variants of perceived instability, presumably, even within the same picture. One variant is gravitational instability, which is disliked and, therefore, reduces the aesthetic appreciation of a picture. The other variant is dynamic instability, which is associated with movement and liked. Because dynamics and movement presumably elicit emotions, which is

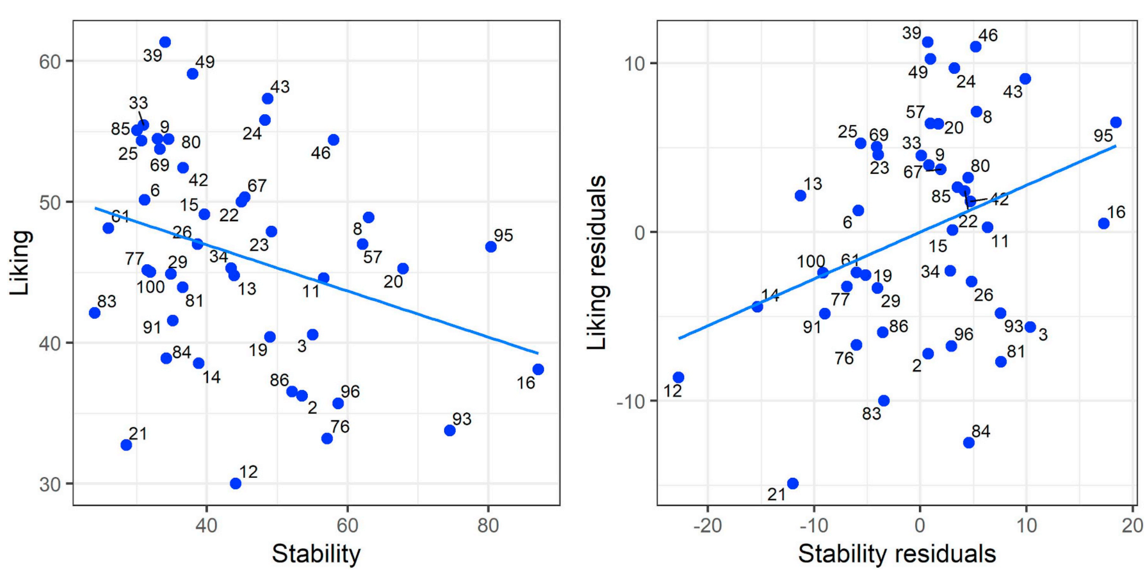


Fig. 2. Left: Relation between stability and liking ratings in Experiment 2. Right: Relation between stability and liking after the effect of implied movement has been removed.

supported by the positive correlation between movement and emotionality, dynamic instability is liked. For collecting emotionality ratings participants were asked how strong their emotions are, because the perceived intensity of enjoyment is often understood as the main source of aesthetic emotion (Menninghaus et al., 2019). Although we did not have assessed the valence of emotionality (Lang, Greenwald, Bradley, & Hamm, 1993; Russell, 1980; Watson & Tellegen, 1985), it is likely that the emotions in the present case were generally positive.

4. General discussion

In the present study, we varied two different types of visual instability and investigated how they affect aesthetic appreciation. Previous studies have suggested that instability could be associated with gravitation stability or with movement (Friedenberg, 2012; Hübner & Fillinger, 2019; Pierce, 1896). To show this more systematically, in our first experiment, we used multi-element pictures showing three rectangles and three colored decorative elements whose composition varied in stability (see Table 1). We assumed that this type of variation is related to gravitational stability and hypothesized that stable pictures are preferred. The observed correlations between stability ratings, liking, and emotionality ratings support our hypothesis. The more instable the pictorial composition looked, the more movement it implied. Because movement was negatively correlated with emotionality, less stable compositions were not liked. Presumably, gravitational stability (van der Helm, 2015) is preferred, because we learn through experience that it is important to arrange things in a way that withstands gravitational forces and thereby prevents damage and injuries. Accordingly, art theory considers stability as an aesthetic habit that plays an important role in composing pictures (Liu et al., 2017).

Our results also indicate that, at least for the applied stimuli, the concepts of stability and balance are rather similar, which is also evident by the results that both stability and balance ratings correlated positively with mirror symmetry. This also demonstrates that mirror symmetry is closely related to stability as well as to balance perception, which is not surprising. However, as the DCM scores correlated with the balance ratings but not with the stability ones, we think that, despite their substantial correlation, stability and balance differ in relevant aspects. Moreover, the fact that the balance ratings correlated much less with the DCM scores than with the stability ratings indicates that the balance ratings reflect both mechanical balance as well as gravitational stability.

A large overlap between the concepts of stability and balance has also been found with Japanese calligraphy as stimuli (Fillinger & Hübner, 2018). In this study, however, there was no relation between balance and liking. Moreover, stability was negatively correlated with liking, i.e., unstable calligraphies were preferred. By inspecting the stimuli, we hypothesized that this surprising result must have been due to the implied movement in these stimuli. If instability serves dynamics and implied movement, which is usually liked, then the picture as a whole is also liked.

This hypothesis was tested in our second experiment, where we used artwork by Karl Otto Götz as stimuli, which largely varied in dynamics. Same as in the first experiment, our participants had to rate the pictures with respect to perceived stability, balance, liking, movement, and emotionality. For these pictures, we expected a negative relation between stability and liking depending on movement. This was indeed the case. Stability ratings correlated negatively with the liking ratings, which replicates our finding for calligraphies (Fillinger & Hübner, 2018). Regression analysis showed that implied movement was responsible for this negative relationship. When we removed the variance from the liking ratings that was accounted for by the movement ratings, then the relation between stability and liking was again positive. These results confirm that there are two types of instability: *gravitational* instability, which is not liked, and *dynamic* instability, which is liked. Moreover, both types can be part of the same rating.

The fact that movement was positively correlated with emotionality supports the idea that dynamic instability is liked, because the accompanying implied movement evokes positive emotions (Dubal et al., 2014). Although we did not assess emotional valence, it is highly likely that the emotions evoked by the present dynamic pictures were positive. In any case, this finding supports the notion that liking is a multidimensional concept that also depends on affective processes (Miller & Hübner, 2019).

Finally, the theoretical concepts of balance and mirror symmetry, as reflected by our objective measures, did not play a role in rating the artwork with respect to perceived stability, balance, and movement or liking and emotionality.

In sum, our study shows that there are two types of perceived instability: gravitational instability, which is disliked, and dynamic instability, which is liked. The former type is related to but different from the common concept of mechanical (im)balance. It applies mainly to relatively static multiple-element pictures. Our experience has taught us that instable arrangements can be dangerous. Accordingly, watching an unstable pictorial composition makes us feel uncomfortable. In contrast, the latter type is related to implied movement, which is usually liked, because dynamics evokes positive emotions. Given that our study shed some light on the relation between perceived stability, balance, and liking, it has relevance for both empirical aesthetics and the arts.

Authors' note

The data as well as the stimuli of the experiments are available at https://osf.io/ebdah/.

CRediT authorship contribution statement

Martin G. Fillinger: Conceptualization, Methodology, Validation, Formal analysis, Investigation, Data curation, Writing - original draft, Visualization. **Ronald Hübner:** Conceptualization, Validation, Writing - review & editing, Resources, Supervision, Project administration.

Acknowledgments

The artwork of Karl Otto Götz has been used with the kind permission of the K.O. Götz und Rissa-Stiftung, Germany.

Author contributions

RH and MF conceived the study. MF constructed or selected the stimuli, collected the data, analyzed the data, and wrote a first draft of the paper. RH revised the paper and supervised the study at all stages.

Declaration of competing interest

The authors declare no competing interests.

References

Arnheim, R. (1982). *The power of the center: A study of composition in the visual arts.* Berkeley: University of California Press.
Bar, M., & Neta, M. (2006). Humans prefer curved visual objects. *Psychological Science, 17*(8), 645–648.
Bouleau, C. (1980). *The painter's secret geometry: A study of composition in art.* New York: Hacker Art Books.
Dubal, S., Lerebours, A.-E., Taffou, M., Pelletier, J., Escande, Y., & Knoblauch, K. (2014). A psychophysical exploration of the perception of emotion from abstract art. *Empirical Studies of the Arts, 32*(1), 27–41. https://doi.org/10.2190/EM.32.1.EOV.4.
Fillinger, M. G., & Hübner, R. (2018). Relations between balance, prototypicality, and aesthetic appreciation for Japanese calligraphy. *Empirical Studies of the Arts*. https://doi.org/10.1177/0276237418805656.
Friedenberg, J. (2012). Aesthetic judgment of triangular shape: Compactness and not the golden ratio determines perceived attractiveness. *i-Perception, 3*(3), 163–175. https://doi.org/10.1068/i0484.
Friedenberg, J., Keating, T., & Liby, B. (2012). Judging the Center of Human Figures:

Evidence for dynamic perception. *International Journal of Brain and Cognitive Sciences, 1*(4), 30–38.
Gershoni, S., & Hochstein, S. (2011). Measuring pictorial balance perception at first glance using Japanese calligraphy. *i-Perception, 2*(6), 508–527. https://doi.org/10.1068/i0472aap.
Gómez-Puerto, G., Munar, E., & Nadal, M. (2016). Preference for curvature: A historical and conceptual framework. *Frontiers in Human Neuroscience, 9*(712). https://doi.org/10.3389/fnhum.2015.00712.
Götz, K. O. (1985). *VAST: Visual aesthetic sensitivity test* (4 ed.). Düsseldorf: Concept.
Greiner, B. (2015). Subject pool recruitment procedures: Organizing experiments with ORSEE. *Journal of the Economic Science Association, 1*(1), 114–125. https://doi.org/10.1007/s40881-015-0004-4.
Hübner, R., & Fillinger, M. G. (2016). Comparison of objective measures for predicting perceptual balance and visual aesthetic preference. *Frontiers in Psychology, 7*(335). https://doi.org/10.3389/fpsyg.2016.00335.
Hübner, R., & Fillinger, M. G. (2019). Perceptual balance, stability, and aesthetic appreciation: Their relations depend on the picture type. *i-Perception, 10*(3). https://doi.org/10.1177/2041669519856040.
Jacobsen, T., & Höfel, L. (2002). Aesthetic judgments of novel graphic patterns: Analyses of individual judgments. *Perceptual and Motor Skills, 95*(3), 755–766. https://doi.org/10.2466/pms.2002.95.3.755.
Kandinsky, W. (1926/1979). *Point and line to plane.* New York: Dover.
Lang, P. J., Greenwald, M. K., Bradley, M. M., & Hamm, A. O. (1993). Looking at pictures: Affective, facial, visceral, and behavioral reactions. *Psychophysiology, 30*(3), 261–273. https://doi.org/10.1111/j.1469-8986.1993.tb03352.x.
Leiner, D. J. (2019). SoSci Survey (version 3.1.06). Retrieved from. https://www.soscisurvey.de.
Liu, J., Dong, W., Zhang, X., & Jiang, Z. (2017). Orientation judgment for abstract paintings. *Multimedia Tools and Applications, 76*(1), 1017–1036. https://doi.org/10.1007/s11042-015-3104-5.
Locher, P. J. (2006). Experimental scrutiny of the role of balance in the visual arts. In P. Locher, C. Martindale, & L. Dorfman (Eds.). *New directions in aesthetics, creativity and the arts* (pp. 19–32). Amityville, New York: Baywood Publishing Company, Inc.
Locher, P. J., Stappers, P. J., & Overbeeke, K. (1998). The role of balance as an organizing design principle underlying adults' compositional strategies for creating visual displays. *Acta Psychologica, 99*(2), 141–161. https://doi.org/10.1016/S0001-6918(98)00008-0.
Mather, G., & Sharman, R. J. (2015). Decision-level adaptation in motion perception. *Royal Society Open Science, 2*(12), 150418. https://doi.org/10.1098/rsos.150418.
McManus, I., Edmondson, D., & Rodger, J. (1985). Balance in pictures. *British Journal of Psychology, 76*(3), 311–324.
McManus, I., Stöver, K., & Kim, D. (2011). Arnheim's gestalt theory of visual balance: Examining the compositional structure of art photographs and abstract images. *i-Perception, 2*(6), 615–647. https://doi.org/10.1068/i0445aap.
Menninghaus, W., Wagner, V., Wassiliwizky, E., Schindler, I., Hanich, J., Jacobsen, T., & Koelsch, S. (2019). What are aesthetic emotions? *Psychological Review, 126*(2), 171–195. https://doi.org/10.1037/rev0000135.
Miller, C. A., & Hübner, R. (2019). Two routes to aesthetic preference, one route to aesthetic inference. *Psychology of Aesthetics, Creativity, and the Arts, 14*(2), 237–249. https://doi.org/10.1037/aca0000241.
Osaka, N., Matsuyoshi, D., Ikeda, T., & Osaka, M. (2010). Implied motion because of instability in Hokusai Manga activates the human motion-sensitive extrastriate visual cortex: An fMRI study of the impact of visual art. *NeuroReport, 21*(4), 264–267. https://doi.org/10.1097/WNR.0b013e328335b371.
Pierce, E. (1894). Studies from the Harvard psychological laboratory (II): Aesthetics of simple forms (I) Symmetry. *Psychological Review, 1*(5), 483–495. https://doi.org/10.1037/h0069000.
Pierce, E. (1896). Studies from the Harvard psychological laboratory (V): The aesthetics of simple forms (II) The functions of the elements. *Psychological Review, 3*(3), 270–282. https://doi.org/10.1037/h0073983.
Ross, D. W. (1907). *A theory of pure design: Harmony, balance, rhythm.* Boston, MA: Houghton, Mifflin.
Russell, J. A. (1980). A circumplex model of affect. *Journal of Personality and Social Psychology, 39*(6), 1161–1178. https://doi.org/10.1037/h0077714.
Silvia, P. J., & Barona, C. M. (2009). Do people prefer curved objects? Angularity, expertise, and aesthetic preference. *Empirical Studies of the Arts, 27*(1), 25–42.
Thömmes, K., & Hübner, R. (2018). Instagram likes for architectural photos can be predicted by quantitative balance measures and curvature. *Frontiers in Psychology, 9*(1050). https://doi.org/10.3389/fpsyg.2018.01050.
Tinio, P. P. L., & Leder, H. (2009). Just how stable are stable aesthetic features? Symmetry, complexity, and the jaws of massive familiarization. *Acta Psychologica, 130*(3), 241–250. https://doi.org/10.1016/j.actpsy.2009.01.001.
van der Helm, P. A. (2015). Symmetry perception. In J. Wagemans (Ed.). *Oxford handbook of perceptual organization* (pp. 108–128). Oxford: Oxford University Press.
Watson, D., & Tellegen, A. (1985). Toward a consensual structure of mood. *Psychological Bulletin, 98*(2), 219–235. https://doi.org/10.1037/0033-2909.98.2.219.
Wilson, A., & Chatterjee, A. (2005). The assessment of preference for balance: Introducing a new test. *Empirical Studies of the Arts, 23*(2), 165–180. https://doi.org/10.2190/B1LR-MVF3-F36X-XR64.