Novelty, complexity, and hedonic value[1]

D. E. BERLYNE[2]
UNIVERSITY OF TORONTO

Two experiments, in which Ss were exposed to sequences of colored shapes, investigated effects on ratings of "pleasingness" and "interestingness" of variables that had previously been shown to affect ratings of "novelty." The results indicate, on the whole, that both pleasingness and interestingness increase with novelty. These findings run counter to those of experiments indicating an inverse relation between novelty and verbally expressed preference. Two further experiments examined effects of some variables that might account for this apparent discrepancy. Homogeneous sequences declined in judged "pleasantness" more than sequences in which several stimuli were interspersed, and simple stimuli became less pleasant as they became less novel, while complex stimuli declined less or became more pleasant. The findings are related to hypotheses regarding mechanisms of hedonic value. Two crucial predictions were confirmed in a fifth experiment.

Two previous experiments (Berlyne & Parham, 1968) investigated determinants of subjective novelty and revealed several factors that can cause a simple, nonrepresentational, visual stimulus (a colored patch of irregular shape) to be rated more or less novel. The experiments to be reported in this paper studied effects of some of the same factors on rated "pleasingness" and "interestingness."

Novelty is one of the "collative" stimulus properties whose importance for motivation theory is becoming more and more apparent, especially in such areas of research as exploratory behavior, experimental aesthetics, developmental psychology, and personality study (Berlyne, 1960, 1963a, 1966). There is now a sizable body of experimental literature (e.g., Berlyne, 1963b; Munsinger & Kessen, 1964; Day, 1965; Berlyne, Ogilvie, & Parham, 1968) on verbal judgments applied to visual patterns of differing complexity, another collative property. Determinants of judged complexity have been identified, and the influence of complexity on judgments of pleasingness (liking, preference) and of interestingness have been studied. The extension of this kind of investigation to novelty seems warranted, and two of the experiments to be reported studied how novelty and complexity interact in determining how pleasing a visual pattern will be rated.

The present state of knowledge with regard to effects of novelty on preference is rather puzzling. Zajonc (1968) reports some experiments of his own, and reviews a fair number of experiments by others, that indicate an inverse relation between novelty and verbally expressed preference. On the other hand, Cantor (1968) reports an experiment with children in which verbally expressed preference increased with novelty. Certainly, novelty tends in many conditions to increase the probability or duration of self-exposure to a stimulus object or manipulation of it, which many would regard as a manifestation of preference or liking. In everyday life, there are times when something is more attractive than something else because it is more familiar, just as there are times when something is found particularly appealing because it is novel. Some resolution of these apparently divergent findings is obviously called for and was one of the objectives of the present investigation.

EXPERIMENTS I AND II

Subjects

Eighty Ss took part in Experiment I and 56 in Experiment II. All were undergraduates taking elementary psychology courses at the University of Toronto.

Stimulus Material

In both experiments, colored patches of irregular shape on a white background were presented in six classes of sequences, and every S went through two sequences belonging to different classes in turn. The sequences were selected from those used in the two experiments reported in the earlier article (Berlyne & Parham, 1968), and incorporated variables that those experiments showed to have significant effects on judged novelty.

The sequence classes can be represented as in Table 1. In Sequence Classes 1 and 2, X differed from Y in color only, whereas X differed from Y in both color and shape in Sequence Classes 1′ and 2′. Stimuli denoted by different letters differed in both color and shape in Sequence Classes 3 and 4.

Experimental Design (see Table 2)

The Ss of each experiment were divided into eight equal groups, containing as far as possible equal numbers of males and females. The key to the notation can be found in Tables 1 and 3 of the previous article (Berlyne & Parham, 1968). Groups 5 and 6 underwent the prefamiliarization (F) treatment, whereas the remaining groups did not. Prefamiliarization consisted of exposing Ss, at the outset of the experiment, to a slide showing all the stimuli to be used in the experiment and asking them to look at the slide until they felt "reasonably well acquainted with the range of pictures shown here" and able to recognize them if they saw them again.

Procedure

Ss of Experiment I were first required to learn the following numerical scale: 7—very pleasing, 6—quite pleasing, 5—slightly pleasing, 4—neither pleasing nor displeasing, 3—slightly displeasing, 2—quite displeasing, 1—very displeasing. For Ss in Experiment II, the words "interesting" and

Table 1
Representation of Sequence Class*

1 and 1′	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	$\underline{X_1}$	Y_{13}	Y_{14}	$\underline{X_L}$
2 and 2′	Y_1	Y_2	$\underline{X_1}$	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	$\underline{X_L}$
3	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7	Y_8	Y_9	Y_{10}	Y_{11}	Y_{12}	Y_{13}	Y_{14}	Y_{15}	$\underline{X_L}$
4	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	$\underline{X_L}$

Subscripts attached to the same letter represent successive appearances of the same stimulus

Table 2
Experiments I and II: Sequences*

Group	Phase	Sequence Class	Sequence															
1	I	(1)	R3	R3	R3	R3	R3	R3	R3	R3	R3	R3	R3	R3	Y3	R3	R3	Y3
	II	(2′)	G2	G2	B1	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	B1
2	I	(2′)	G2	G2	Y3	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	Y3
	II	(1)	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	B1	R1	R1	B1
3	I	(1′)	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	Y3	G2	G2	Y3
	II	(2)	R1	R1	B1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	R1	B1
4	I	(2)	R3	R3	Y3	R3	R3	R3	R3	R3	R3	R3	R3	R3	R3	R3	R3	Y3
	II	(1′)	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	G2	B1	G2	G2	B1
5 (F) and 7 (NF)	I	(3)	1	1	1	1	1	1	1	1	1	1	1	1	1	1	1	Y3
	II	(4)	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	B1
6 (F) and 8 (NF)	I	(4)	1	4	5	6	7	8	9	10	11	12	13	14	15	16	17	Y3
	II	(3)	2	2	2	2	2	2	2	2	2	2	2	2	2	2	2	B1

** For key, see Tables 1 and 3, Berlyne & Parham, 1968*
F = with prefamiliarization; NF = without prefamiliarization

"uninteresting" were substituted for "pleasing" and "displeasing." There was then a test for ability to supply the number corresponding to each judgment and vice versa. After both S and E were satisfied that the scale had been mastered, the room was darkened, and the appropriate two sequences of stimuli were presented in turn. Each stimulus was projected for 9 sec on the screen. A blank slide intervened for 9 sec between the two phases, while E announced, "Another series will follow immediately after this." During the exposure of each stimulus, S had to assign a degree of pleasingness (Experiment I) or interestingness (Experiment II) to it and utter the corresponding number.

Results

The curves in Fig. 1 show the results for the different sequence classes, and Table 3 sums up the statistical analyses. Included are orthogonal comparisons corresponding to variables that were found to affect subjective novelty in the previous investigation, analyses of linear components of declining trends over monotonous portions of sequences, and sign tests showing how judged pleasingness or interestingness rose when a new stimulus appeared after a monotonous run and sank again when the stimulus figuring in the monotonous run reappeared.

Among the pleasingness data (Experiment I), two important interactions involving variables mentioned in Table 3 were found. First, the ratings of the final stimulus, X_L, were subject to a significant interaction between presence or absence of prefamiliarization (F vs NF) and the homogeneity or heterogeneity of the sequence preceding X_L (Sequence Classes 3 vs 4): $F(1,36) = 4.97$, $p < .05$. The mean for Condition 4F (3.0) was markedly lower than the means for the other three combinations of treatments, which were close together at 4.8-5.0. Secondly, there was a significant interaction between phases and the prefamiliarization variable with regard to the linear component of the trend over the first 15 stimuli in Sequence Class 3: $F(1,504) = 5.75$, $p < .025$. The decline was markedly less steep for the NF treatment in Phase II than in the other three combinations.

In the trend analyses, the method appropriate for equally spaced values of the independent variable was used. They were equally spaced numerically, but the question of whether they were equally spaced psychologically may be raised. In this connection, it is worth noting Osgood, Suci, and Tannenbaum's (1957, p. 327) report of "... fairly satisfying evidence that 7-step scales, defined by the linguistic quantifiers, 'extremely,' 'quite' and 'slightly,' do yield nearly equal psychological units in the process of judgment."

Discussion

It will be seen that not all the variables that were found to influence judged novelty significantly in the previous investigation had significant effects on judged pleasingness or interestingness. There was even a hint that interestingness may reach a maximum with a certain degree of familiarity in the unexpected finding of a significantly higher mean rating (Experiment II) in Ss who had had the prefamiliarization treatment than in those who had not (X_L, Sequences 3 and 4) had this treatment. Nevertheless, the results tend strongly to favor one of the two divergent views mentioned earlier over the other: Both pleasingness and interestingness appear to increase with novelty. There is a marked general resemblance between the graphs in Fig. 1 and those representing mean novelty ratings that were presented in the previous article (Berlyne & Parham, 1968). In particular, pleasingness and interestingness, like subjective novelty, decline with prolonged repetition of a particular stimulus, and they both rise temporarily when a monotonous sequence is interrupted by a single appearance of a contrasting stimulus.

In experiments on complexity (Berlyne, 1963; Day, 1965, 1967, 1968), conditions have been found in which relatively high pleasingness coincides with relatively low interestingness, namely when complexity is particularly low. But as far as novelty is concerned, nothing of this sort seems to occur. Pleasingness and interestingness seem generally to go together, although the curves in Fig. 1 suggest that interestingness may decline faster.

EXPERIMENT III

Aims

The results of Experiments I and II, like those of Cantor's (1968) experiment, indicate a direct relation between novelty and hedonic value, whereas the experiments reported and reviewed by Zajonc (1968) suggest an inverse relation. So we must look for differences in material or procedure that might account for these discrepant outcomes.

Three possibilities suggest themselves, and some attempt was made in Experiment III to investigate all of them:

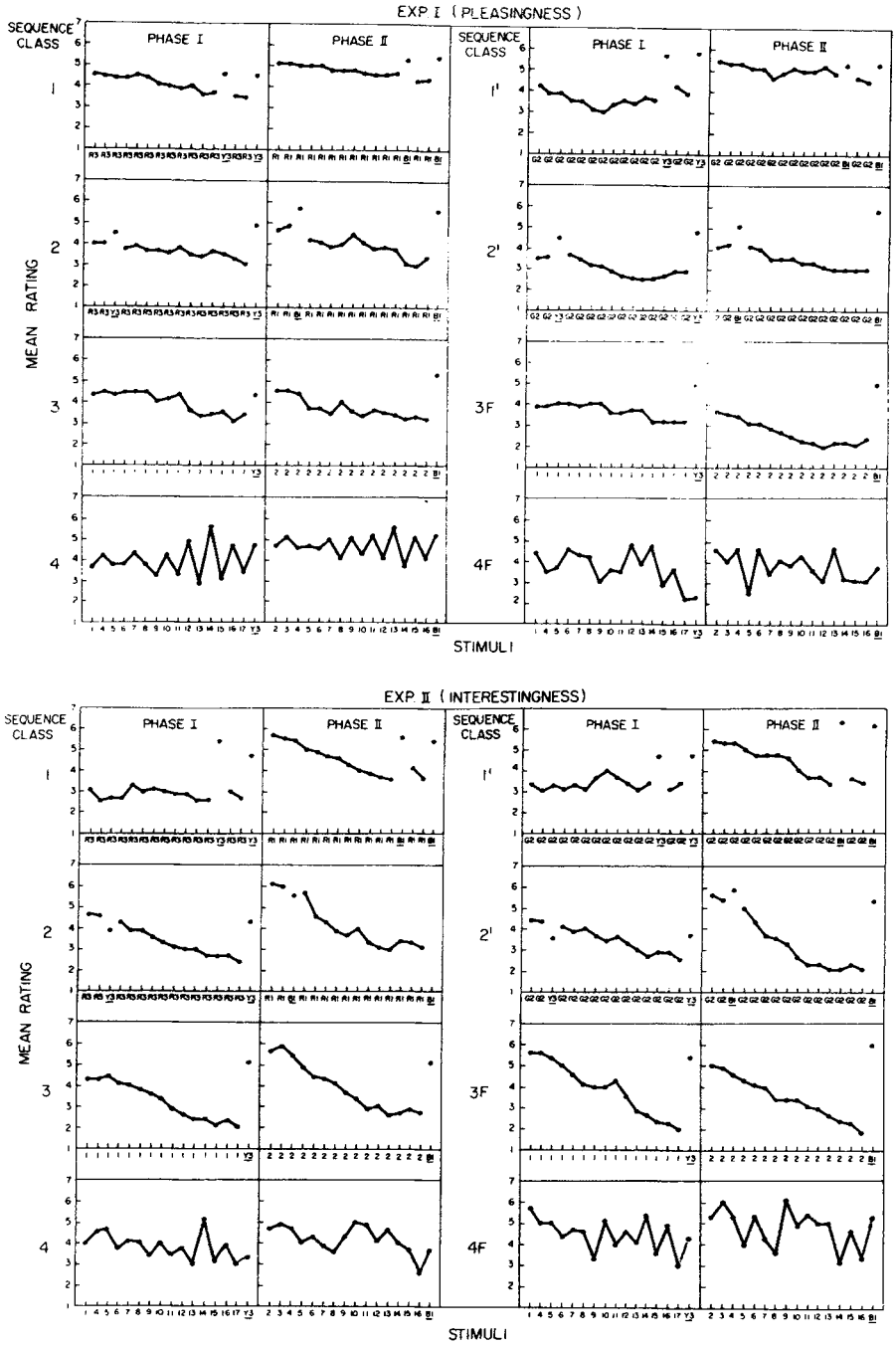

Fig. 1. Experiments I and II: Mean ratings.

(1) The stimuli used in Experiments I and II, like the black-and-white patterns from the Welsh Figure Preference Test used by Cantor, were rather simple, whereas those used in the experiments performed and cited by Zajonc (words, paralogs, Chinese ideograms, photographs of faces) were rather complex. If stimuli sometimes become less appealing and sometimes become more appealing as they become familiar, this could be explained by the interaction of two antagonistic processes, of which sometimes one prevails and sometimes the other. As a stimulus is repeated or prolonged, its appeal may gradually succumb to a "tedium" factor. Preference may, on the other hand, rise with repetition or prolongation through the action of a "positive-habituation" factor. It may reflect increased liking for a pattern as information is absorbed from it and it becomes assimilated or organized. One would expect the tedium factor to be relatively stronger when stimuli are simple and low in information content, whereas there will be more scope for the information processing presumed to underlie the positive-habituation factor when stimuli are complex. Experiment III, therefore, used less complex and more complex stimuli in order to see if the hedonic value of the former would decline, and that of the latter would rise, as they lost their novelty.

(2) The stimuli used in the experiments supporting a direct relation between novelty and hedonic value were meaningless, whereas those used in the experiments supporting inverse relation were meaningful and symbolic, being verbal or pictorial. In Experiment III, both nonrepresentationl patterns and reproductions of representational paintings were used as stimulus material. They cannot, of course, be regarded as representative samples of representational and nonrepresentational visual patterns, so that generalizations from any comparison between their effects are hazardous. The intention was simply to see whether different or even opposite trends might appear with the two kinds of stimuli.

(3) In Experiments I and II (but not in Cantor's experiment), one stimulus was presented repeatedly 12 or 15 times. In Zajonc's experiments, different stimuli were exposed 1, 2, 5, 10, and 25 times before Ss rated them, and the stimuli receiving different numbers of preexposures were intermingled in the preliminary sequence. One would expect monotonous repetition to strengthen the tedium factor, while the appearance of several stimuli in a sequence might well mitigate this factor and allow the positive-habituation factor to gain the upper hand. So, in Experiment III, every S went through both a homogeneous sequence, comprising 10 presentations of one and the same stimulus, and a heterogeneous sequence of 40 presentations, in which 10 presentations of each of four stimuli were interspersed.

Subjects

Forty-eight students, for the most part not students of psychology or of fine arts, were recruited in the corridors of the University of Paris-Nanterre. They participated in the experiment in parties of two to six.

Stimulus Material

Four nonrepresentational (N) patterns were taken from a set (Berlyne, 1963b) that has been used over the years in a number of experiments studying the effects of complexity variables on

Table 3
Experiments I and II

A. Orthogonal Dyadic Comparisons

Variable	Comparison	Pleasingness (Experiment I) Means	F	df	p	Interestingness (Experiment II) Means	F	df	p
First or Second Appearance of X	X_1 (1, 1′, 2, 2′) vs X_L (1, 1′, 2, 2′)	5.1 / 5.3	–	–	NS	4.9 / 5.1	–	–	NS
Number of Ys Before First X	X_1 (1, 1′) vs X_1 (2, 2′)	5.2 / 5.0	–	–	NS	5.4 / 4.7	–	–	NS
Recency of Previous X	X_L (1, 1′) vs X_L (2, 2′)	5.3 / 5.3	–	–	NS	5.3 / 4.6	4.23	1,24	~0.5
Degree of Change from Y to X	X_L (1, 2) vs X_L (1′, 2′)	5.1 / 5.4	–	–	NS	4.9 / 5.0	–	–	NS
Homogeneous or Heterogeneous Sequence Before X	X_L (3F, 3NF) vs X_L (4F, 4NF)	4.9 / 4.0	7.02	1,36	<.025	5.4 / 4.1	17.76	1,24	<.001
Prefamiliarization	X_L (3F, 4F) vs X_L (3NF, 4NF)	4.0 / 4.9	4.97	1,36	<.05	5.2 / 4.3	4.48	1,24	<.05

B. Linear Components of Trends

Stimuli	Comparison	Pleasingness (Experiment I) F	df	p	Interestingness (Experiment II) F	df	p
First 12 of (1) and (1′)	Overall Decline	31.09	1, 418	<.001	63.10	1,286	<.001
First 15 of (3F) and (3NF)	Overall Decline	195.04	1, 504	<.001	556.39	1,336	<.001
First 15 of (4F) and (4NF)	Overall Decline	21.22	1, 504	<.001	28.29	1,336	<.001
First 15	Steeper Decline in (3F, 3NF) than in (4F, 4NF)	5.13	1,1008	<.025	85.60	1,672	<.001

C. Sign Tests

Sequence	Comparison	Pleasingness (Experiment I) Distribution (Without Ties)	p	Interestingness (Experiment II) Distribution (Without Ties)	p
(1, 1′)	X_1 > Preceding Y	27/8	<.01	23/3	<.01
(1, 1′)	X_1 > Succeeding Y	22/8	<.05	24/1	<.01
(1, 1′)	X_L > Preceding Y	30/5	<.01	24/1	<.01
(2, 2′)	X_1 > Preceding Y	22/7	<.01	13/12	NS
(2, 2′)	X_1 > Succeeding Y	24/7	<.01	14/7	NS
(2, 2′)	X_L > Preceding Y	35/1	<.01	25/2	<.01
(3F, 3NF)	X_L > Preceding Y	23/10	<.05	26/1	<.01

exploratory behavior and other processes. Two simple (NS) items and two complex (NC) items represented, respectively, the lower and upper reaches of the subjective-complexity continuum, as scaling experiments (Day, 1965; Berlyne, Ogilvie, & Parham, 1968) have confirmed. The four representational (R) patterns were black-and-white reproductions of paintings. The representational-simple (RS) items were portraits of single figures on a plain background, and the representational-complex (RC) items were crowded canvases replete with a multitude of human figures and other details.

One NS, one NC, one RS, and one RC item were selected by tosses of a coin to be designated by the number 1 and to be elements of the homogeneous sequences. They consisted of less complex members of the third pair in Category C and the more complex member of the first pair in Category XC (Figs. 1 and 2, Berlyne, 1963b), Ramsay's *Norman McLeod*, and Rubens's *La Kermesse*. The remaining items were designated by the number 2 and used for the heterogeneous sequence. They consisted of the less complex member of the fourth pair in Category A, the more complex member of the second pair in Category XA, Raeburn's *Portrait of a Man* and Rubens's *Massacre of the Innocents*.

Experimental Design and Procedure

Ss were divided into eight groups of six. Groups 1-4 went through a homogeneous sequence followed by the heterogeneous sequence, while Groups 5-8 had the heterogeneous sequence before the homogeneous sequence. The homogeneous sequence comprised 10 successive presentations of Item NS1 for Groups 1 and 5, of NC1 for Groups 2 and 6, of RS1 for Groups 3 and 7, and of RC1 for Groups 4 and 8. The heterogeneous sequence was the same for all groups. It consisted of 10 presentations each of Items NS2, NC2, RS2, and RC2, randomly intermingled with the restriction that every item appear once in each consecutive set of four.

Each presentation consisted of a 4-sec projection on a screen, and there was a 4-sec interval between consecutive presentations. An interval of 1 min 30 sec intervened between the two phases, and there was a break of 1 min between the first and the second halves of the heterogeneous sequence to allow the slide magazine to be changed.

S had in front of him a booklet, each page of which bore a horizontal line divided into seven compartments, with the word "Désagréable" (unpleasant) to its left and the word "Agréable" (pleasant) on the right. The scale thus resembled those used for the semantic differential (Osgood, Suci, & Tannenbaum, 1957). After each presentation, S had to mark one of the seven compartments of the appropriate scale. He had been told that the scale went from "extremely unpleasant" to "extremely pleasant" and that the center represented the point of indifference "that is to say, neither pleasant nor unpleasant."

Results

Separate analyses of variance were carried out on the ratings for the homogeneous sequence and heterogeneous sequence. The results are shown in Fig. 2.

Homogeneous sequence. The N items were rated higher than the R items–means = 3.7, 2.8, $F(1,40) = 6.11$, $p < .025$. This unexpected finding is peripheral to the main objectives of the experiment. It is presumably connected with the fact that contemporary taste in painting, particularly in intellectual circles, favors the nonrepresentational. And such forms of representational painting as are now in

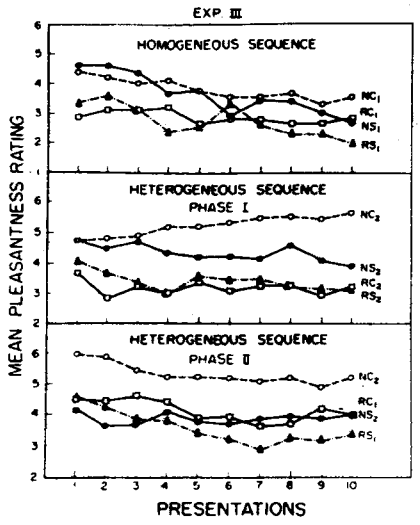

Fig. 2. Experiment III: Mean ratings.

vogue (e.g., Pop Art, *Nouvelle Figuration*) are certainly very different from our Baroque and Neoclassical R items.

A trend analysis showed the linear component of the overall declining trend to be significant: $F(1,432) = 52.51$, $p < .001$. The linear component was steeper for N items than for R items–$F(1,432) = 4.49$, $p < .05$ and steeper for S items than for C items–$F(1,432) = 11.61$, $p < .005$. The declines were, however, significant for the N items alone ($F = 42.94$, $p < .001$), the R items alone ($F = 13.03$, $p < .001$), the S items alone ($F = 55.81$, $p < .001$), and the C items alone ($F = 7.25$, $p < .01$).

None of the comparisons or interactions involving the phase difference (whether the homogeneous series occurred first or second) reached significance.

Heterogeneous sequence. In the heterogeneous sequence, the overall mean rating was once again higher for N items than for R items–means = 4.23, 3.24; $F(1,1728) = 481.44$, $p < .001$–and higher for C items than for S items–means = 4.03, 3.44; $F = 173.82$, $p < .001$. There was, however, a significant interaction between these two dichotomies: $F = 97.32$, $p < .001$. It turned out that the NC item was rated significantly higher than the NS item–means = 4.7, 3.7; $F = 26.56$, $p < .001$–but the means for the RC and RS items (3.3, 3.2) were close together.

Turning to the trend analysis, we find both the overall linear component–$F(1,1794) = 31.21$, $p < .001$–and the quadratic component–$F = 9.29$, $p < .005$–to be significant. The linear component was significantly steeper for the S items than for the C items–$F = 10.60$, $p < .005$. It was also steeper in Phase II (i.e., in groups having the heterogeneous sequence second)–$F = 15.35$, $p < .001$.

There was, however, a significant triple interaction involving the R/N, S/C, and phase variables–$F = 12.91$, $p < .001$. For this reason, the results for the two phases are plotted separately in Fig. 2. Inspection of the curves shows that the triple interaction resulted from the greater dissimilarity of trends in Phase I than in Phase II. So the data for this phase were subjected to a separate trend analysis. In Groups 5-8, which had the heterogeneous sequence first, the overall linear component was not significant. The interaction between the linear component and the R/N and S/C variables was significant–$F(1,864) = 9.65$, $p < .005$. Data for the R and N items were therefore examined separately. The linear components for the RS and RC items did not differ significantly, and there was a significant decline over both R items taken together–$F = 7.70$, $p < .01$. Linear components differed significantly for the NS and NC items–$F = 29.22$, $p < .001$. The NS item produced a significant decline –$F = 10.01$, $p < .005$, while the NC items produced a *rise* with a significant linear component–$F = 19.20$, $p < .001$.

Discussion

Our hypothesis predicting a decrease in the hedonic value of simpler patterns and an increase in that of more complex patterns as repetition reduces novelty receives partial confirmation. This is precisely what happened with the N patterns of the heterogeneous sequence in Phase I. And in both sequences, there was a significantly steeper decline with the S stimuli than with the C stimuli. Nevertheless, ratings of both S and C stimuli declined in the heterogeneous sequence when it came second and in the homogeneous sequence. We must, therefore, suppose that the homogeneity-heterogeneity variable played some part and that monotony favors the tedium factor. Furthermore, Ss were apparently more susceptible to the tedium factor during the second phase of the experiment. This is shown by the fact that ratings declined for all stimuli of the heterogeneous sequence in Phase II, including the NC stimulus that produced a rise in Phase I.

The N/R variable also proved to have some influence, but it took the opposite form to that suggested by the observation that experiments indicating an inverse relation between novelty and hedonic value have used meaningful, symbolic stimuli, while those indicating a direct relation have used meaningless stimuli. In both the homogeneous and the heterogeneous sequences, the representational stimuli made for a steeper decline than the nonrepresentational stimuli.

EXPERIMENT IV

Aims

The results of the three experiments reported so far certainly cast doubt on the validity for all kinds of stimulus material of Zajonc's conclusion that hedonic value increases with familiarity. There is, however, one more factor that must be examined. In Zajonc's experiments, S had simply to look at the stimulus patterns during the familiarization phase and did not have to record judgments until a later phase. In our experiments, a rating had to be made after every presentation. This might very well have made the procedure more laborious and therefore more tedious. If no ratings had been required during familiarization, it is possible that positive habituation would have come to the fore, producing a rise in preference even for the simpler stimuli as they became more familiar. Experiment IV was carried out to check this possibility.

The necessity for a judgment after every representation might have affected the ratings by making Ss pay more attention to the stimuli. To see if this was an operative factor, half of the Ss in Experiment IV were given instructions designed to induce close attention, while the other half were not, but repeated judgments were required of neither group during familiarization.

Subjects

Forty-eight students were recruited as in Experiment III. They all went through the heterogeneous sequence alone and had to rate stimuli only during the first four presentations of the sequence (during which each of the four stimuli appeared once) and during the last four presentations (during which again each stimulus appeared once). After the first four presentations had been concluded, 24 Ss, comprising the attention-instruction (AI) group, were told that they would see the pictures again a number of times and that they should look at them closely because they would later be questioned about them. The next 32 presentations (8 of each stimulus) then followed, and Ss were then instructed about the second set of ratings just before the last four presentations were given. The 24 Ss of the no-attention instruction (NAI) group were simply told after the first four presentations that they would see the stimuli again a number of times.

Results

Neither the main effect nor any of the interactions involving the presence or

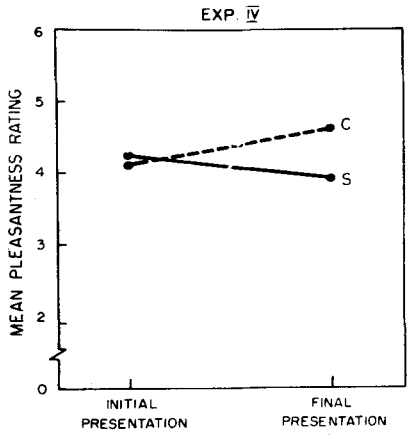

Fig. 3. Experiment IV: Mean ratings.

absence of attention instructions reached significance. So the results of the AI and NAI groups will be reported together. The main effects of the initial-rating/final-rating and N/R and S/C dichotomies were not significant. The only significant comparisons in the analysis-of-variance table were the N/R by S/C interaction—$F(1,322) = 51.65$, $p < .001$—and the comparison pertaining to the primary aim of the experiment, namely the Initial-Final by S/C interaction—$F(1,322) = 5.84$, $p < .025$.

The former interaction reflects the fact that the RS stimulus was rated higher than the RC stimulus—$F(1,322) = 15.53$, $p < .001$—whereas the NC stimulus was rated higher than the NS stimulus—$F = 38.90$, $p < .001$. As for the second of these interactions, the mean final rating of the C stimuli was significantly higher than their initial rating—$F(1,322) = 4.49$, $p < .05$—while, in the case of the S patterns, the final rating was lower than the initial rating but not significantly so (see Fig. 3).

Finally, the results of Experiment IV were examined in conjunction with those for the heterogeneous sequence in Groups 5, 6, 7, and 8 of Experiment III, i.e., the groups that had the heterogeneous sequence first. The double Experiments by S/C interaction and the triple Experiments by Initial/Final by S/C interaction were far from significant. Over both experiments, the Initial-Final by S/C interaction was significant—$F(1,483) = 10.42$, $p < .005$. The mean final rating for the S stimuli (3.3) was significantly lower than the mean initial rating for those stimuli (4.2)—$F(1,483) = 6.00$, $p < .025$. For the C stimuli, on the other hand, the mean final rating (4.6) was significantly higher than the mean initial rating (4.1)—$F = 6.19$, $p < .025$.

Furthermore, the triple Initial-Final by S/C by N/R interaction and the quadruple interaction involving these variables and the experiments variable are of interest in view of the different results that were obtained with N and R patterns when the heterogeneous sequence came first in Experiment III. However, neither of these interactions reached significance.

GENERAL DISCUSSION

Our findings support our hypothesis that the hedonic value of complex stimuli tends to rise as they become less novel while the opposite holds for simple stimuli. We have, however, some evidence that monotonous, homogeneous sequences are more conducive than varied sequences to a decline in hedonic value after familiarization.

In our experiments, the representational-nonrepresentational variable influenced ratings but in ways that are rather complicated and may very well be peculiar to the specific kinds of stimulus used. So it would be hazardous to draw any general conclusions from these effects, except that we have no support for the conjecture stated at the beginning of this article that a rise in hedonic value with familiarization is characteristic of meaningful symbolic material, while a decline is characteristic of nonsense material. We have also no indication that changes in hedonic value take a different course when Ss simply look at stimuli as they appear from when they have to record judgments during familiarization or look with special attentiveness.

It was suggested earlier in this article that the apparent discrepancies among experimental findings might be resolved if we supposed that changes in hedonic value depended on the interaction of two antagonistic factors, a tedium factor that might plausibly preponderate when a simple stimulus was encountered repeatedly and a positive-habituation factor that might preponderate during repetition of a complex stimulus.

It may be worth our while to consider these two processes a little further and, in particular, to relate them to a tentative view of the mechanisms underlying hedonic value for which the evidence has been reviewed elsewhere (Berlyne, 1967). According to this view, hedonic value (a term meant provisionally to cover both reward value, as judged by the capacity of a stimulus to reinforce an instrumental response, and preference or pleasure, which is reflected in verbal evaluations) can be produced in either of two ways. First, any condition that drives arousal up to a high level is aversive, so that a stimulus that subsequently reduces arousal will be rewarding and pleasant. Secondly, any stimulus condition that produces a moderate increase in arousal will be rewarding and pleasant. The manner in which the second or arousal-increase mechanism works can be summed up by a reinterpretation of the Wundt curve (Fig. 4). Here, the abscissa is taken to represent "arousal potential," a term meant to cover all the stimulus properties that tend to raise arousal, including novelty and complexity. Positive hedonic value reaches a maximum with moderate arousal potential (stimulation producing a moderate arousal increment) and then, as arousal potential increases further, hedonic value takes on lower and lower positive values and finally becomes negative.

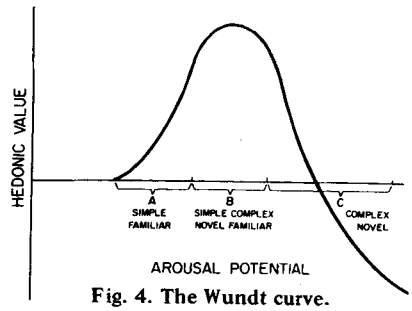

Fig. 4. The Wundt curve.

A high degree of novelty means a high degree of arousal potential, so that, as a stimulus becomes familiar and loses its novelty, we must imagine ourselves moving along the horizontal axis of the Wundt curve from right to left.

The decreasing appeal of a simple stimulus with familiarization is easy to account for in terms of the curve. A stimulus that is high in novelty but low in complexity will have medium arousal potential (somewhere around Region B in Fig. 4). Progressive loss of novelty through repeated or prolonged exposure would correspond to a leftward movement along the horizontal axis from Region B to Region A accompanied by falling hedonic value. An additional mechanism is, however, apt to intervene eventually. Protracted exposure to stimulation of low information content apparently leads to an increase in at least the noncerebral indices of arousal and to negative hedonic value (Berlyne, 1960; Schultz, 1965). A boredom factor, conceived in this way, could explain why ratings of simple stimuli sometimes end up negative.

The positive-habituation factor that seems to preponderate with complex stimuli in at least some circumstances could have at least two explanations, and it is difficult at this stage to know the relative importance of either. First, the initial impact of a complex stimulus pattern can be expected to engender uncertainty, conflict, disorientation (Berlyne, 1960, 1963a, 1966, in press). Opportunities for

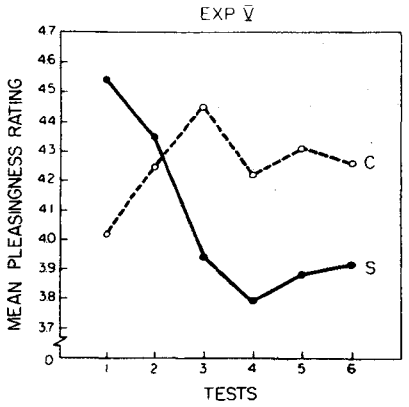

Fig. 5. Experiment V: Mean ratings.

further acquaintance with the pattern will provide scope for absorption of additional information and for the perceptual and ideational processing that enables uncertainty and conflict to be resolved as elements are discriminated, classified, recognized, and grouped together as sub-wholes. This can evidently be a source of pleasure, presumably dependent on the arousal-reduction mechanism. However, the arousal-increase mechanism to which the Wundt curve is relevant is also likely to come into play. A stimulus that is high in both novelty and complexity will be high in arousal value and thus correspond to a point on the horizontal axis in Region C. As it lost its novelty, the point representing its arousal potential would shift towards Region B accompanied by an increase in hedonic value.

Skaife's (1967) findings with auditory stimulus material are consonant with ours and open to a similar interpretation. She used musical sequences varying in what she called "complexity," but, since it involved varying degrees of deviation from normal melodies, it might better be designated "surprisingness" or "uncertainty." Whatever name is applied to it, it can be recognized as a constituent of arousal potential. When repeated over 20 days, low-"complexity" sequences that were initially rated pleasant gradually lost their pleasantness, whereas high-"complexity" sequences that were initially given low ratings underwent a significant rise in judged pleasantness.

Our interpetation in terms of the Wundt curve implies that, if familiarization had proceeded further, ratings of the simple patterns would have continued to decline, whereas those of the complex patterns would have climbed to a peak and then dropped. Skaife (1967) recorded such a rise followed by a fall with some of her musical sequences and some of her Ss. So did Alpert (1953) with repetitions of an unfamiliar rhythmic sound pattern. It seemed desirable, however, to verify these predictions with our visual material. Experiment V was undertaken for this purpose.

EXPERIMENT V

Subjects

Forty undergraduates from elementary psychology classes at the University of Toronto took part, two at a time.

Procedure

The procedure was the same as for Experiment IV, except that it was prolonged. S went through six tests, during each of which the four patterns were presented in a random order and had to be rated on a 7-point Osgood scale from "displeasing" to "pleasing." Between each test and the next, the four patterns were each presented eight times in the random order that had been used between the initial and final presentations in Experiment IV, but no judgments were recorded.

Results

The overall mean was higher for the R patterns (4.4) than for the N patterns (3.8): $F(1,897) = 30.12$, $p < .001$. This was opposite to the preference recorded in Experiment III. This difference might have reflected differences in taste between French and Canadian students. It might, on the other hand, have been due to the fact that the projector used in Toronto (a Kodak Carousel) produced brighter images than the projector used in Paris (a Liesegang nonautomatic), since the N slides consisted mostly of white background. There was also a significant interaction between the N/R and C/S dichotomies–$F(1,897) = 15.26$, $p < .001$. The means for the SN, SR, CN, and CR patterns were, respectively, 3.5, 4.6, 4.0, and 4.2.

To turn to the comparisons that are germane to the chief purpose of the experiment, the Tests by C/S interaction was significant–$F(5,897) = 2.65$, $p < .025$. But neither the Tests by N/R nor the Tests by C/S by N/R interaction approached significance. Consequently, curves for the two S patterns and the two C patterns are presented in Fig. 5.

Ferguson's (1965) nonparametric trend analysis was used. Since the hypotheses under investigation implied a decline for S patterns and a rise followed by a decline for C patterns, one-tailed p values were relevant. With regard to C patterns, there was significant bitonicity–$z = 2.19$, $p < .015$. The monotonic and tritonic components were not significant. As for the S patterns, there was both a significant monotonic component–$z = 4.30$, $p < .001$–and a significant bitonic component–$z = 2.83$, $p < .005$.

Discussion

The two predictions from the interpretation in terms of the Wundt curve thus received confirmation, and this explanation of the positive-habituation and tedium factors is therefore tenable. The C curve reached a maximum at the third test. After that, it dropped slowly and erratically. This may have been because rewarding effects of perceptual processing came into play. The bitonic component in the predominantly downward trend of the S curve seems to represent a flattening out that is compatible with the shape of the Wundt curve (see Fig. 4 and Berlyne, 1967), if we make the assumption that arousal potential drops linearly as number of presentations increases. The rise in the S curve over the last three tests was far from significant–$z = 0.52$. Since the S curve sank only slightly below the indifference point (4.0), boredom seems to have played no more than a minor role.

REFERENCES

ALPERT, R. Perceptual determinants of affect. Unpublished MA thesis, Wesleyan University, 1953.

BERLYNE, D. E. *Conflict, arousal and curiosity*. New York: McGraw-Hill, 1960.

BERLYNE, D. E. Motivational problems raised by exploratory and epistemic behavior. In S. Koch (Ed.), *Psychology–A study of a science*. Vol. 5. New York: McGraw-Hill, 1963a. Pp. 284-364.

BERLYNE, D. E. Complexity and incongruity variables as determinants of exploratory choice and evaluative ratings. Canadian Journal of Psychology, 1963b, 17, 274-290.

BERLYNE, D. E. Curiosity and exploration. Science, 1966, 153, 25-33.

BERLYNE, D. E. Arousal and reinforcement. In D. Levine (Ed.), *Nebraska Symposium on Motivation, 1967*. Lincoln, Nebr: University of Nebraska Press, 1967. Pp. 1-110.

BERLYNE, D. E. *Psychobiology and aesthetics*. New York: Appleton-Century-Crofts, in press.

BERLYNE, D. E., OGILVIE, J. C., & PARHAM, L. C. C. The dimensionality of visual complexity, interestingness and pleasingness. Canadian Journal of Psychology, 1968, 22, 376-387.

BERLYNE, D. E., & PARHAM, L. C. C. Determinants of subjective novelty. Perception & Psychophysics, 1968, 3, 415-423.

CANTOR, G. N. Children's "like-dislike" ratings of familiarized and unfamiliarized visual stimuli. Journal of Experimental Child Psychology, 1968, 6, 651-657.

DAY, H. Exploratory behavior as a function of individual differences and level of arousal. Unpublished PhD thesis, University of Toronto, 1965.

DAY, H. Evaluations of subjective complexity, pleasingness and interestingness for a series of random polygons varying in complexity. Perception & Psychophysics, 1967, 2, 281-286.

DAY, H. The importance of symmetry and complexity in the evaluation of complexity, interest and pleasingness. Psychonomic Science, 1968, 10, 339-340.

FERGUSON, G. A. *Nonparametric trend analysis.* Montreal: McGill University Press, 1965.
MUNSINGER, H., & KESSEN, W. Uncertainty, structure and preference. Psychological Monographs, 1964, 78, No. 9.
OSGOOD, C. E., SUCI, G. J., & TANNENBAUM, P. H. *The measurement of meaning.* Urbana: University of Illinois Press, 1957.
SCHULTZ, D. P. *Sensory restriction: Effects on behavior.* New York: Academic Press, 1965.
SKAIFE, A. M. The role of complexity and deviation in changing musical taste. Unpublished PhD thesis, University of Oregon, 1967.
ZAJONC, R. B. Attitudinal effects of mere exposure. Journal of Personality & Social Psychology Monograph Supplement, 1968, 9, No. 2, Part 2.

NOTES

1. This investigation was supported by Research Grant APB-73 from the National Research Council of Canada. Experiments III and IV were carried out while the author held a Visiting Professorship at the Institut d'Esthétique et des Sciences de l'Art, University of Paris, awarded jointly by the Scientific Affairs Division of NATO and the Minna-James-Heineman Foundation, while on sabbatical leave from the University of Toronto. The data of Experiments I and II were collected by Jerry Colglazier, those of Experiments III and IV by Lydie Boudet, and those of Experiment V by Mary Louise King and Kathryn Davies. The author is indebted to them for their able collaboration. Experiments III and IV were carried out in the Laboratory of Experimental Psychology of the University of Paris-Nanterre, thanks to the cooperation of Professor R. Francès.

2. Address: Department of Psychology, University of Toronto, Toronto, Ontario, Canada.

(Accepted for publication January 15, 1970.)