JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR VOLUME 5, NUMBER 3 JULY, 1962

INSTRUCTIONS AND STIMULUS CATEGORIZING IN A MEASURE OF STIMULUS GENERALIZATION[1]

DAVID R. THOMAS AND KEVIN MITCHELL[2]

KENT STATE UNIVERSITY

In Experiment I, three groups of 20 *Ss* each were exposed to a light of 550 mμ (yellowish-green) for 60 sec and then viewed a random sequence of wavelengths with instructions to respond *only* to the original color. The instructions given the three groups were worded differently in an attempt to vary the strength of a set-to-discriminate assumed to be created by this procedure. The three groups produced similar gradients, each with a peak of responding at 540 mμ, in agreement with Kalish's (1958) published gradient for the 550 mμ standard stimulus value. It was suggested that the nature of the task is such that a strong discriminatory set is produced regardless of the wording of the instructions.

A temporal analysis of the gradient as it develops during the testing revealed that initially the peak of responding occurs at 550 mμ; but as testing progresses, it shifts gradually in the direction of the shorter wavelengths (purer greens). Experiment II was performed to test the generality of the phenomenon of regression to the primary color. Two groups of 20 *Ss* each were tested for generalization following exposure to 510 mμ (bluish-green) and 525 mμ (pure green), respectively. We predicted that the 510 mμ gradient would reveal a progressive shift toward the longer wavelengths (purer greens), whereas the 525 mμ gradient would show no tendency to shift. The results were strikingly in accord with these predictions.

We concluded that although a physiological process could not be ruled out, the verbal labeling of the standard stimulus value may well be responsible for the regression of the gradient toward the primary color.

In 1958, Kalish reported a study which introduced a new and highly efficient procedure for the study of stimulus generalization in human *Ss*. A brief outline of his procedure follows. Human *Ss* are exposed to a monochromatic stimulus for 60 sec. They are instructed to remember the color and are told that they will subsequently be required to identify it. The *Ss* are then presented with a random sequence of stimuli, including the original, with instructions to respond (*i.e.*, to lift a finger from a telegraph key) *only* to the original color. Each stimulus is presented an equal number of times. Because the trials are discrete, a measure of response probability to each stimulus is obtained. This measure typically demonstrates a gradient in response strength; *i.e.*, *Ss* respond to novel stimuli to the degree that they are physically similar to the original.

The principal difference between the Kalish procedure for the study of generalization and most of the others is that the tendency to respond to the standard stimulus value is created by instructions rather than by training. These instructions explicitly prohibit responding to novel stimuli, and generalization is said to occur to the degree that *Ss* fail to follow the instructions.

Because the instructions to the *Ss* play so central a role in this procedure, differently worded instructions might produce very different results. By prohibiting responding to novel stimuli, Kalish's instructions probably induce a set-to-discriminate in his *Ss*. Our hypothesis was that instructions worded so as to strengthen or to weaken this set would be reflected in steeper or flatter generalization gradients, respectively. Experiment I was performed to test this hypothesis.

EXPERIMENT I

Method

Subjects. The *Ss* were 60 undergraduate women enrolled in introductory psychology courses at Kent State University. All *Ss* had normal color vision, as determined with the Dvorine color perception test (Dvorine, 1944).

Apparatus. A Skinner-type, key-pecking apparatus, modified for use with human *Ss*,

[1]This research was supported in part by U.S.P.H.S., N.I.M.H. Research Grant M-4203.

[2]Now at Michigan State University.

was used. The box was approximately 15 in. long, 14 in. wide, and 11 in. high, and was painted flat black. The front wall of the box was made of transparent Plexiglas. The *S* sat in a chair 2 1/2 ft from the key, which was a circular plastic disc 7/8 in. in diameter. A Cambridge Thermionic Corp. monochromator, Model B Ser. 1066, equipped with an Olympus Model 201250 6-v, 5-amp light source, illuminated the key. The patch of color on the key was approximately 4 ml in luminance. The only other light in the room was a 7.5-watt "night light" on *E*'s side of a black cloth screen separating *S* from *E*.

A telegraph key was used as the manipulandum. This key was placed on the table next to the experimental box within easy reach of S's right hand. It was wired so that its release would activate a signal light on *E*'s side of the screen, thus signifying a response. The box was equipped with an electrically operated shutter which interrupted the monochromator beam when *E* threw a switch.

Procedure. The *S*s were divided unsystematically into three groups of 20 *S*s each. Each *S* received one of the three sets of instructions. In the reproduction of the instructions which follows, the set used by Kalish is labeled "neutral"; the instructions designed to strengthen the tendency to discriminate are designated "discriminatory-set instructions"; and those designed to weaken the discriminatory tendency are called "generalization-set instructions."

1. Neutral instructions

"This is an experiment in color preception. At the beginning of the experiment a specific color will be presented through the small hole in front of you. Try to keep this color in mind because you will be asked to identify it later. After one (1) minute this color will be turned off and you will place your finger and press down on the telegraph key in front of you. I will give the signal 'ready' and a few seconds later a color will again be presented. You must decide whether this is the original color shown you at the start of the experiment. If it is, *lift* your hand as rapidly as you can from the key. It it is not, keep pressing on the key.

"I will say the word 'ready' whenever I am about to present a color and you should be pressing the key at that time. We are going to try some practice trials. Now we are going to run through a series exactly as we could do if this were the real experiment."

At this point, *E* exposes 600 mμ for 60 sec and then presents 610 mμ, 590 mμ, 600 mμ, 620 mμ, and 580 mμ.

"Now we are going to begin the experiment. Remember, try to keep the original color in mind and respond as rapidly as you can, lifting your finger only when the original color appears. Do not be disturbed, however, if you should respond to other colors. Any questions?"

2. Discriminatory-set instructions

These instructions were the same as those above except that the following paragraph was added:

"Specifically, your task will be to make the fewest possible errors. An error is defined as a response to a color that is not the original color. *In other words, you must not respond to any color other than the original color.* Your accuracy score will be compared with that of others in your class."

The discriminatory-set instructions ended with the reminder, "*Remember, you must not respond to any color other than the original color.*"

3. Generalization-set instructions

These instructions were the same as the neutral ones except that the following paragraph was added:

"Specifically, your task will be to make the fewest possible errors. An error is defined as a failure to respond to the original color. *In other words, you must respond when the original color is presented again.* Your accuracy score will be compared with that of others in your class"

The generalization-set instructions ended with the reminder, "*Remember, you must respond to the original color when it is presented again.*"

After the reading of the instructions and the performance of the practice series, the

actual experiment was begun. The standard stimulus (550 mμ) was exposed for 60 sec. Then, 6 randomized series of 8 test stimuli plus the standard stimulus were presented. The stimuli were 510 mμ through 590 mμ in 10-mμ steps, inclusive. Each test stimulus was presented for a period of 3 sec and was followed by a blackout of 5 to 10 sec, allowing sufficient time for *E* to record the presence or absence of a response and to change the monochromator setting.

Results and Discussion

Figure 1 presents the data for the total number of responses the *Ss* of the three experimental groups made to the different stimuli. Clearly, the results fail to support the experimental hypothesis. The three gradients are strikingly similar to each other and to Kalish's published gradient for this stimulus value. Either the different sets of instructions failed to induce different sets, or the sets were induced but they failed to affect the generalization gradients. The former possibility seems more likely. After reading the instructions, most *Ss* spontaneously recited a simplified condensation of them. These spontaneous recitations typically included only the most basic aspects of procedure, omitting the set-inducing parts. Thus, although the complete written instructions to the three experimental groups did differ, the more informal self-instructions apparently did not. The three sets of instructions used in this study were purposely made as similar as possible except for the presumably set-inducing parts. As a result, those aspects of the instructions the different groups shared in common may have had more power than the differences between them. Although a different wording of the instructions might have had greater influence on the *S*'s behavior, the nature of the experimental task was such that strong discriminatory-set probably would have developed in any case. Indeed, when human *Ss* are used, a discriminatory-set can probably be avoided only if no mention is made of the stimulus dimension over which

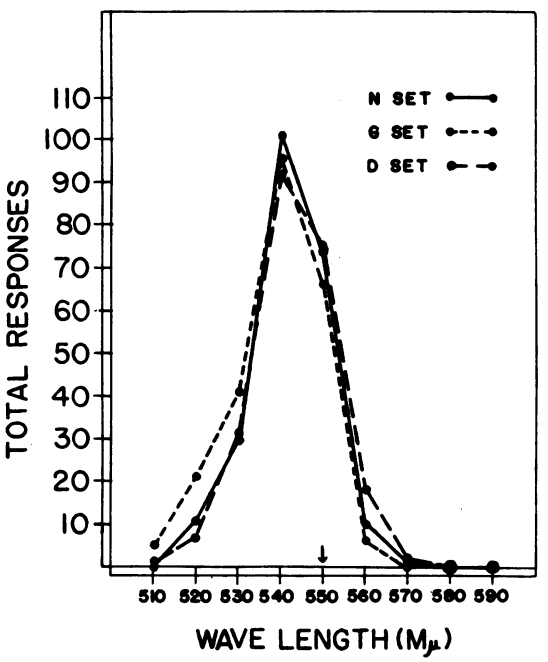

Fig. 1. Generalization gradients of the three experimental groups.

Table 1
The Generalization Gradients for the 20 *Ss* in the N-set Group

Ss	*Wavelength (mμ)*								
	510	*520*	*530*	*540*	*550*	*560*	*570*	*580*	*590*
1			1	6	1				
2		3	5	6	4	1			
3		1	1	6	6				
4		1		4	6	5			
5		4	5	5	2		1		
6			6	5	2				
7				5	2				
8			1	6	4				
9		1	3	5	5				
10				6	2				
11				6	5				
12				4	5				
13				4	1				
14				3	6				
15				6	5				
16				2	6	4			
17			1	6	4				
18				6	3				
19			4	6	5				
20		1	3	5					
	Σ	11	30	102	74	10	1		

generalization is to be measured. A conditioning procedure is one situation in which this end might be accomplished. Research is currently underway in our laboratory to compare generalization following instructions and various conditioning procedures.

A digression might be valuable here to indicate the degree to which the group curves in Fig. 1 are representative of individual performance. Table 1 presents the data which went into the plotting of the generalization gradient of the N-set group in Fig. 1. Although the intersubject variability is considerable, the group picture is clearly not biased. Of the 20 *Ss*, 14 responded with greater frequency to 540 mμ than to 550 mμ. The steepness of the group gradient is also typical since 15 of the 20 *Ss* responded to only two or three different stimulus values.

In all three experimental groups studied here, and in Kalish's (1958) experiment as well, the peak of the gradient around a standard stimulus value of 550 mμ was 540 mμ. The consistency with which this displacement of the peak occurs suggests that more than chance is involved. It is interesting to note whether this displacement appears at the start of the generalization testing session or develops gradually during the session. Consequently, the data for the three experimental groups were pooled, and gradients were plotted separately for the first series of stimulus presentations; Series 1 and 2 combined; Series 3 and 4 combined; and Series 5 and 6 combined.

Figure 2 shows clearly that the location of the gradient shifts gradually; it moves toward the shorter wavelengths as the test progresses. The peak of responding is located at 550 mμ only for the first two series of test stimuli.

The measure used in Fig. 2 is the total number of responses per stimulus. Such a measure is susceptible to possible distortion by extreme scores of a few deviant individuals. A more conservative measure, therefore, is the frequency of *Ss* with a peak of responding at a given stimulus value. Figure 3 (and Fig. 5 and 7 to follow) uses this measure of the number of *Ss* who show a peak of responding at the different wavelength values. Thus, in this figure, if all *Ss* had shown greatest response strength to the original stimulus, the value of the ordinate at that point would be equal to the number of *Ss*. If a given *S* responded

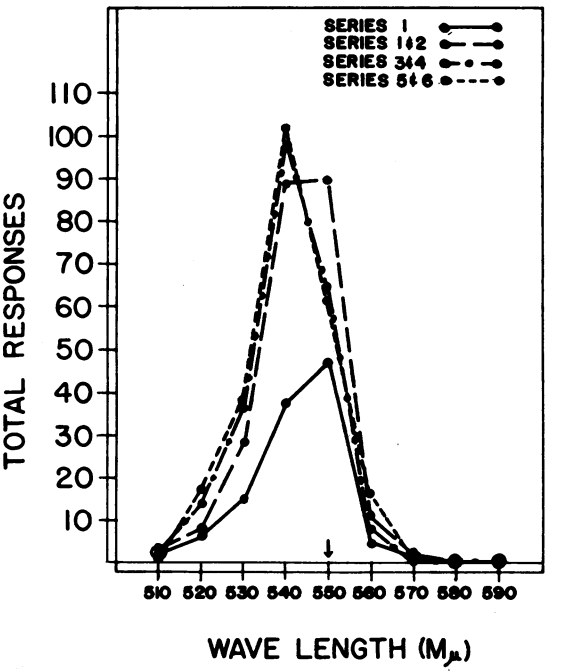

Fig. 2. Generalization gradients during different stages of testing, in terms of total responses per stimulus.

equally frequently to two or more stimuli, the score of 1 for that individual was equally divided among the stimuli involved. Therefore, the total of all ordinate values remained equal to the number of *Ss* in these curves.

Figure 3 presents the gradients for the first series of stimulus presentations, Series 1 and 2 combined, Series 3 and 4 combined, and Series 5 and 6 combined, plotted in terms of this frequency measure. The striking similarity of the temporal analyses of the gradient using the two different measures demonstrates the reliability of the gradual-displacement phenomenon for this standard stimulus value.

The value 550 mμ appears as a yellowish-green, whereas 540 mμ is a purer green. Thus, the progression of the *Ss*' stimulus preference involves a regression toward the primary color. Before becoming concerned with an explanation of this phenomenon, its generality should first be evaluated by the use of additional standard stimulus values. If generalization were tested around a standard stimulus value seen as bluish-green (*e.g.*, 510 mμ), a shift of the peak toward the longer wavelengths (purer greens) would be expected to occur during testing. On the other hand, a gradient obtained with a standard stimulus value seen as pure green (*e.g.*, 525 mμ) should show no

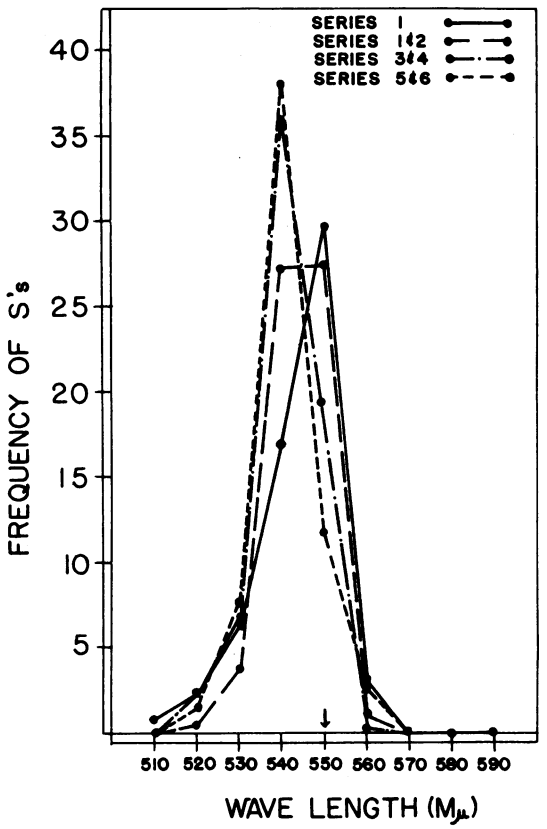

Fig. 3. Generalization gradients during different stages of testing, in terms of the number of *Ss* with peak responding to the different test stimuli.

tendency to shift at all. Experiment II was performed to test these hypotheses.

EXPERIMENT II

Method

Subjects. The *Ss* were 40 undergraduate women enrolled in introductory psychology courses at Kent State University. All *Ss* had normal color vision, as determined with the Dvorine color perception test (Dvorine, 1944).

Apparatus. The apparatus was the same as that used in Experiment I.

Procedure. The procedure was the same as that used in Experiment I with certain exceptions. The *Ss* were divided unsystematically into two groups of 20 *Ss* each. All *Ss* received the same instructions, the neutral instructions from Experiment I. One group of *Ss* viewed 510 mμ (bluish-green) as the standard stimulus and was subsequently tested for generalization with test stimuli from 470 mμ through 550 mμ in 10-mμ steps, inclusive. The second group viewed 525 mμ (pure green) and then was tested with 485 to 565 mμ in 10-mμ steps, inclusive. The generalization testing was carried out in the same manner as in Experiment I.

Results and Discussion

Figure 4 presents the total response gradients for the group exposed to a standard stimulus value of 510 mμ, taken from test Series 1, Series 1 and 2 combined, Series 3 and 4 combined,

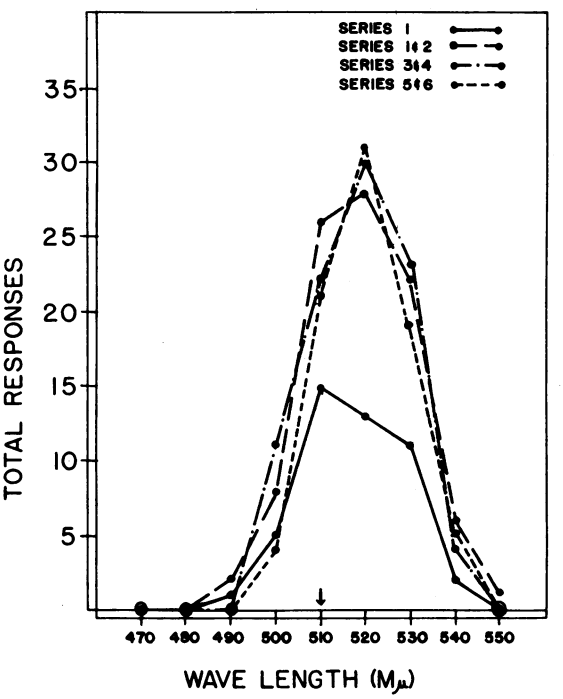

Fig. 4. Generalization gradients during different stages of testing, in terms of total responses per stimulus.

and Series 5 and 6 combined. Figure 5 shows a comparable analysis in terms of the number of *Ss* who showed peaks of responding at the different test stimuli. Both analyses clearly indicate that the peak of response strength progresses gradually toward the longer wavelengths, *i.e.*, the purer greens.

Figure 6 presents the total response gradients for Series 1, Series 1 and 2 combined, Series 3 and 4 combined, and Series 5 and 6 combined, for the group exposed to the standard stimulus value of 525 mμ. Figure 7 presents a parallel analysis in terms of the number of *Ss* with a peak of responding at different test stimuli. Both analyses agree that the generalization gradient following exposure to 525 mμ does

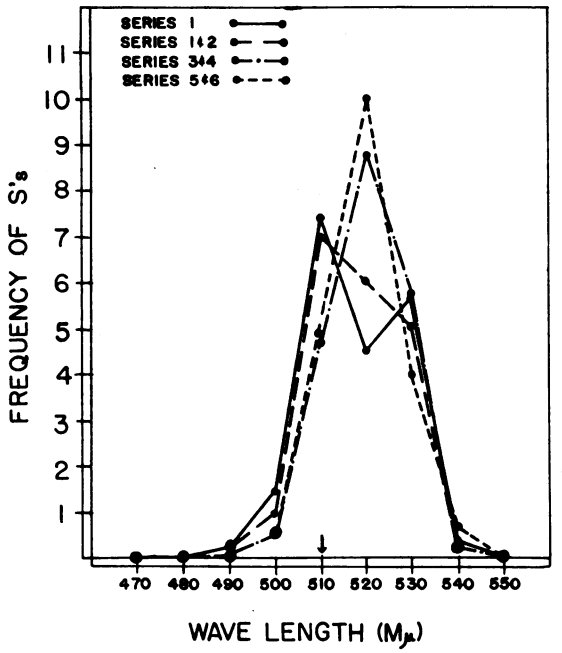

Fig. 5. Generalization gradients during different stages of testing, in terms of the number of *Ss* with peak responding to the different test stimuli.

not tend to shift its location along the wavelength dimension during the testing session. The results of Experiment II are thus in clear support of the experimental hypotheses.

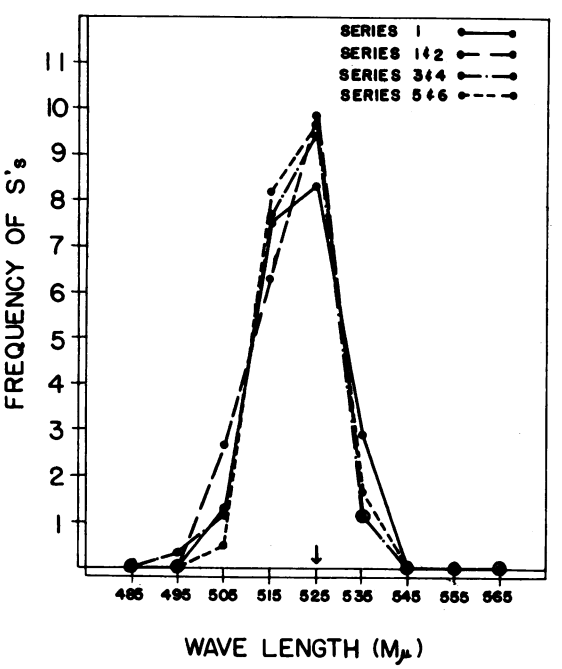

Fig. 6. Generalization gradients during different stages of testing, in terms of total responses per stimulus.

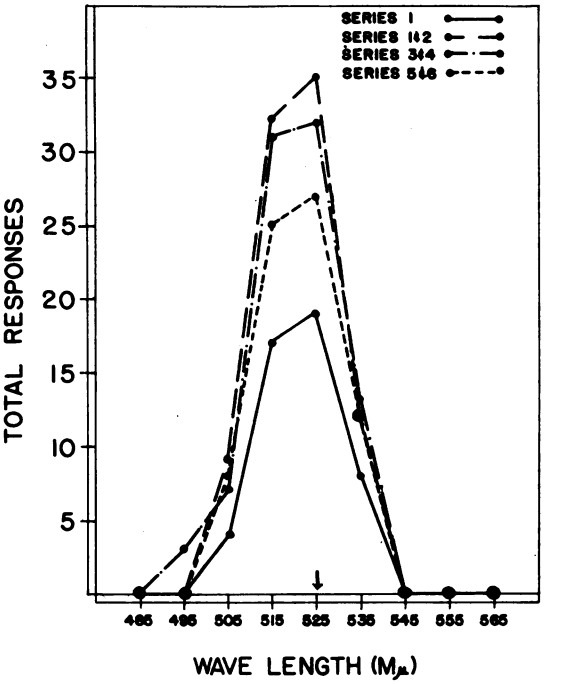

Fig. 7. Generalization gradients during different stages of testing, in terms of the number of *Ss* with peak responding to the different test stimuli.

Consequently, a regression to the primary color does occur, and it is reflected in the gradient of stimulus generalization as measured here. It is interesting to speculate on what kind of mechanism may be responsible for this effect. Both 550 mμ (yellowish-green) and 510 mμ (bluish-green) are labeled green. Indeed, human *Ss* have a history of conceptualizing a wide range of wavelengths as green, blue, red, etc. It is tempting to assume that this labeling of the standard stimulus value distorted its retention much in the same manner as the labels distorted the retention of the drawings in the classical Carmichael, Hogan, and Walter (1932) experiment. This is a reasonable hypothesis, but our data unfortunately do not provide a test of it.

The alternative possibility exists that a physiological process rather than a psychological one is basic to the displacement effect, and that color naming may have nothing whatsoever to do with it. Studies of generalization using pigeon *Ss* and an operant-conditioning technique (*e.g.*, Guttman & Kalish, 1956; Thomas & King, 1959) find no evidence for a regression to the primary color. Of course, the subject-related and procedure-

related differences between the human and the pigeon studies are so great as the question of the validity of this comparison. At present, the question of the relative merits of a psychological *vs.* a physiological interpretation of the regression effect reported in this study must remain an open one.

REFERENCES

Carmichael, L., Hogan, H. F., and Walter, A. A. An experimental study of the effect of language on the reproduction of visually perceived form. *J. exp. Psychol.*, 1932, **15**, 73-86.

Dvorine, I. *Dvorine color perception testing charts.* Baltimore, Md.: Waverly Press, 1944.

Guttman, N., and Kalish, H. I. Discriminability and stimulus generalization. *J. exp. Psychol.*, 1956, **51**, 79-88.

Kalish, H. I. The relationship between discriminability and generalization: A re-evaluation. *J. exp. Psychol.*, 1958, **55**, 637-644.

Thomas, D. R., and King, R. A. Stimulus generalization as a function of level of motivation. *J. exp. Psychol.*, 1959, **57**, 323-328.

Received October 20, 1961