

Relative olfactory intensity perception as mediated by ratio-range category scale responses¹

R. A. M. GREGSON,² M. J. MITCHELL, M. B. SIMMONDS, AND J. ELISABETH WELLS
UNIVERSITY OF CANTERBURY, NEW ZEALAND

Scales whose categories are labeled with ranges of ratio values are compared with verbal category scales and magnitude estimation. Relative perceived intensities of Eugenol odor were scaled by power law methods, using 102 Ss, five scaling methods—one verbal, two numerical, and two magnitude estimation—and making comparisons against two alternate odor reference standards. Variations in the psychophysical exponent values derived under each condition were examined. Comparisons between scale types were made: numerical ratio-range category scales may behave as magnitude estimations or as category scales depending on the way responses are scored by the experimenter.

A comparative study of three methods of scaling relative olfactory intensity is reported. The scaling methods used were magnitude estimation, with and without a standard initially given by E, category scaling with verbally labeled steps, and a form of category scale with numerically labeled steps that we have called a Ratio Range Category Scale, hereafter abbreviated to RRCS.

If the steps in a category scale mediating relative judgments are given integer numerical designations, n , two choices are immediately obvious; either the steps are labeled with single numbers 'n' so that with reference to a standard m they are relatively 'n/m' ($n > m$ or $n = m$ or $n < m$), or each step is a bounded mutually exclusive range of such numbers, say, 'n₁/m to n₂/m.' This latter is an RRCS. Following Stevens, m may be put as 10 or 100 for convenience.

The behavior of an RRCS may empirically transpire to be like a verbal category scale or like a magnitude estimation task, because it has features that it shares with each of the previously employed response modes. It is known from a number of studies (Galanter & Messick, 1961; Helm, Messick, & Tucker, 1961; Svenson, 1967) that there is frequently a curvilinear relation between category and magnitude estimation procedures on prothetic continua, of which olfactory intensity is supposed to be an example. Hence, if a scaling based on an RRCS is linear on a corresponding verbally labeled category scale and curvilinear on a magnitude estimation procedure then the category step construction of the RRCS is playing a dominant role, whereas if the RRCS yields a linear relationship on magnitude estimation but a nonlinear one on a category scale then the numerical labeling of the RRCS is dominant over the category step construction in determining Ss' response processes, when the two features are in conflict.

An RRCS may also be analyzed both as an equal interval scale, by arbitrarily numbering the categories by integers in sequential order, and as a scale whose intervals are as defined by its numerical labeling, which is not necessarily equal-interval, on the same data. The relation between the scalings derived from these two methods of analysis provides further data for comparison with the previously reported curvilinear relation between magnitude estimation and category scaling.

There is a loss of information in moving from magnitude estimation to RRCS arising from the increased error variance consequent upon regressing all responses in the range ($n_1/m - n_2/m$) as though they were at the mean response ($(n_1 + n_2)/2m$). This will slightly but not seriously diminish the variance taken up by a regression analysis of responses on stimuli

if the regression is in fact linear under some transformation.

Helm, Messick, and Tucker (1961) have advanced a model in which category scales differ from magnitude estimation because the latter mediate differences in sensation whereas the magnitude scale mediates corresponding ratios. The model agrees with Eisler's interpretation of the problem (1962a, b, 1963) and predicts the logarithmic regression of category values upon magnitude estimations, with a small correction introduced by Eisler that is apparently necessitated by departures from Weber's law. This model provides another way of determining whether RRCS behave more like category scales or more like a magnitude estimation task.

The manner in which Ss employ an RRCS may also be predicted from Parducci's range-frequency models (1963, 1965). Here, Ss are postulated to divide a range of stimuli into proportionate subranges, and to use each subrange in proportion to the frequency of stimuli falling in that range. From these postulates Parducci deduces that middle categories of a scale cover approximately the same psychological differences in terms of discriminability.

The olfactory stimulus, Eugenol, was here varied over a concentration range of about 40 to 1. Previous investigators (Engen & Pfaffman, 1959, 1960) suggest that the information capacity of this sensory continuum is low (1.5 bits for intensity, 3.0 bits for hedonic tone), and clearly less than the capacity provided in the response modes utilized here. The psychophysical power law exponent for olfactory intensity (Stevens & Galanter, 1957; Jones, 1958a, b; Reese & Stevens, 1960) has been established at about 0.50 in the range 0.40 to 0.60 for a range of substances not including Eugenol. The data presented here provide a further check on these earlier determinations.

METHOD

Subjects

The Ss were 102 undergraduates, mostly first-year psychology students, 51 male and 51 female, all naive with respect to the psychophysical theory involved.

Stimuli

The stimuli were Eugenol, CS_2 : $CH_2CH_2C_6H_5(OCH_3)$: OH, in varying degrees of concentration in filtered air, presented in a continuous stream, in the following concentrations:

Table 1
Eugenol Concentrations Used

Flowmeter setting	Calibrated Abs concn.	Ratios re std	Ratios re std
	-1	4	6
	10 mg/liter air	(x 10)	(x 10)
1	(0.087)	0.7	0.4
2	0.48	3.1	1.7
3	0.99	6.4	3.7
4	1.55	10.0	5.7
5	2.17	13.8	7.8
6	2.75	19.5	10.0
7	3.35	21.3	12.2
8	3.96	25.2	14.2
9	4.82	30.0	17.1

Table 2
Category Scale Step Labeling

(1) Verbal	(2) RRCS1	(3) RRCS2
Much less intense	0 to 2	0 to 1
Less intense	2+ to 4	1+ to 4
Slightly less intense	4+ to 6	4+ to 7
Very slightly less intense	6+ to 8	7+ to 9
Possibly less intense	8+ to 10-	9+ to 10-
Equally intense	Equally intense, 10	Equally intense, 10
Possibly more intense	10+ to 12	10+ to 18
Very slightly more intense	12+ to 16	18+ to 26
Slightly more intense	16+ to 24	26+ to 34
More intense	24+ to 40	34+ to 40
Much more intense	40+ and more intense	40+ and more intense

Notes: (1) Setting 1 gave results that were discarded in the final analysis of power function experiments, as residual effects in the apparatus, detected on calibration checks, made its perceived intensity too high. (2) Eugenol was selected because it is mildly pleasant at most concentrations (Sandusky & Parducci, 1965), is relatively free from trigeminal irritation, and was suitable for use with the method of dilution used in the olfactometer, where a saturated vapor is obtained by bubbling an airstream through a column of glass beads whose interstices are filled with the odorant liquid.

Response Scales

The three forms of category scale used were interchangeable in the experimental setting: to each of the 11 scale steps there corresponded one Morse Key that Ss pressed to give a response.

The labeling of the scale steps for each of the three forms is given in Table 2.

All three category-scale types were accompanied by instructions to the effect that they referred to the intensity of a second stimulus in a pair judged relative to the first in a pair. In the cases of RRCS1 and RRCS2 Ss were instructed that the first stimulus always had an intensity of 10 units.

The magnitude estimation (ME) procedures followed as closely as possible the paradigms given in Harper and Stevens (1964). In the anchored instructions (ME1) Ss were presented with a standard stimulus concentration, told it was 10 units, and were thereafter free to pair any number to any stimulus presentation. In the unanchored instructions (ME2) Ss were free to give any stimulus any numerical value. The standard was included in the stimulus series and randomly presented (but not specified as being a standard) during ME2 as often as in the paired comparison, ME1, and category scale procedures, to control for adaptation effects.

There were thus 10 conditions in all: verbal category, RRCS1, RRCS2, ME1, and ME2, each with two standard stimulus concentrations, 1.55 and 2.75, $\times 10^{-1}$ mg/l. In nine of these conditions 10 Ss were used; in the last ME2, with 2.75 standard, 12 Ss were available.

Procedure

The experiment was conducted in the Olfactory Psychophysics Laboratory at the University of Canterbury, Ss being tested individually over a single session of about 45 min. S was seated in the cubicle of an olfactometer built on the continuous-flow bypass circuit principle originally devised by Ough and Stone (1961), and modified slightly so that S was enclosed in a controlled and continually replenished atmosphere of clean air. The olfactometer allows continuous and infinitely variable concentration odor presentation in an airstream ejected from a grille on the end of a glass tube suspended just in front of S's nose. Odor concentration was read off from calibrated Fischer and Porter flowmeters, and air temperature, line pressure, and

odorant temperature recorded for calibration corrections as necessary. All responses were made to a time cycle signaled to S by E during the experiment, and S made all responses by pressing keys, or in the magnitude estimations by speaking into a microphone. The key-pressing responses and their latencies were recorded on an Esterline Angus multichannel pen recorder.

The values of the standard stimulus (corresponding to the median category on the three scale types, and always presented first in the paired-comparison tasks used) were Scale Steps 4 and 6 of Table 1, that is, 1.55 and 2.75×10^{-1} mg/l. The choice of standard affects the physical concentration ratios presented in the manner shown in the last two columns of Table 1, their skewness also implies that Ss are forced to make a different relative frequency pattern of usage of the scale categories (at least in the RRCS1 and RRCS2 cases) because the proportion of the stimuli that are below the median category differs over the two standards, and the actual ratios involved, particularly in the suprmedian subrange, impose different restrictions on response distributions.

RESULTS

The properties of the RRCS relative to the other scale types here investigated are established by examining the functional relations between the various intensity continua obtained by using the scales or transformations based upon them. This is conveniently done in four parts:

- (1) Methods of transformation to numerical scales.
- (2) and (3) Interscale relations and scale-stimulus relations.
- (4) Psychophysical power function exponents.

(1) The verbal scale can be transformed into a numerical scale by using Parducci's range-frequency theory in which steps are assumed to be equivalent to an equal-ratio partitioning of the total range of the stimulus continuum available, and new equivalent values for the steps are then substituted. From these the mean ratings of each of the stimuli can then be obtained directly. The necessary transformation assumes that b , the psychophysical exponent, is equal to unity, and then the actual b involved can subsequently be read off directly from the resultant plot of log stimulus values against log transformed scale values. The results of using Parducci's theorem to transform the verbal scales are set out in Table 3 and are used again in Table 6.

The mean values from the RRCSs are calculated directly by assuming that the numerical labeling of the steps can be taken at its face value, as in a magnitude estimation task. Comparison is then possible.

Table 3
Transformation of Verbal Category Scale Steps to Equal Ratio Scale Steps, on the Equal Ratio Partitioning Assumption Derived from Parducci's Model, with Exponent Set at $b = 1.0$. Log Values and Corresponding Absolute Ratio Values are Given.

Step	Ratios re std (expressed in 10^{-1} mg/liter)		Log ratios re std (expressed in 10^{-1} mg/liter)	
	std at 1.55	std at 2.75	std at 1.55	std at 2.75
1	3.1	1.7	.491	.230
2	3.9	2.4	.593	.384
3	5.0	3.5	.695	.538
4	6.3	4.9	.796	.692
5	7.9	7.0	.898	.846
6	10.0	10.0	1.000	1.000
7	12.5	11.1	1.095	1.047
8	15.5	12.4	1.191	1.093
9	19.3	13.8	1.286	1.140
10	24.1	15.4	1.382	1.186
11	30.0	17.1	1.477	1.233

The lowest stimulus intensity is taken as 0.375×10^{-1} mg/liter, the highest as 4.82×10^{-1} mg/liter. This ignores stimulus 1, which is about equal to stimulus 2.

Table 4
Logarithms of Mean Perceived Relative Stimulus Intensities for Each Scale Type and Standard Stimulus Concentration, for Stimuli on Flowmeter Settings 2 through 9, as given in Table 1. Flowmeter Setting 1 is Omitted as Unreliable.

Flowmeter setting stimulus conc.	Verbal Scale ‡		RRCS1		RRCS2		ME1		ME2	
	standard stimulus		standard stimulus		standard stimulus		standard stimulus		standard stimulus	
	1.55	2.75	1.55	2.75	1.55	2.75	1.55	2.75	1.55	2.75
2	0.740	0.533	0.598	0.542	0.550	0.519	0.504	0.130	0.316	0.096
3	0.829	0.650	0.826	0.710	0.798	0.688	0.682	0.516	0.561	0.423
4 ^w	0.938	0.777	0.963	0.891	0.986	0.880	0.827	0.751	0.796	0.603
5	1.116	0.856	1.081	0.906	1.144	0.969	0.989	0.917	0.791	0.667
6 ^x	1.088	0.894	1.084	1.005	1.208	1.067	1.030	0.937	0.850	0.808
7	1.165	0.895	1.119	1.016	1.260	1.096	1.048	1.002	0.927	0.862
8	1.210	1.001	1.211	1.094	1.287	1.249	1.127	1.017	0.938	0.868
9	1.279	1.047	1.218	1.155	1.375	1.334	1.186	1.126	1.026	1.002

‡ The verbal scale values are arrived at using the transformation of Table 3.

^w The setting corresponds to standard 1.55×10^{-1} mg/liter.

^x This setting corresponds to standard 2.75×10^{-1} mg/liter.

(2) The correspondence between the stimulus values based on the transformed verbal scale and the directly interpreted RRCS values are set out in Table 4. Table 4 also lists for comparison the stimulus values generated by the more conventional magnitude estimation procedures, since the category scaling procedures were undertaken in order to see what relation obtains between these various methods. If the transformations that are applied to verbal scales and the RRCSs are in some way equivalent to a magnitude estimation procedure, then we should obtain at least a linear relation between each pair-wise comparison of the three scalings for each standard stimulus value. For the case where the standard is 1.55, plotting all functions against the anchored ME1 case as abscissa, the ME2 case is linear with the same slope but less in value, the verbal scale and the two RRCSs have slightly different slopes and are all consistently greater in value, and RRCS2 has the steepest slope. All the relationships can be taken as linear. For the case where the standard is 2.75, again plotting all functions against the anchored ME1 as abscissa, the ME2 is slightly less steep and consistently less in value, the trend being linear. The three other scales are again greater in their corresponding values, but now the trends are nonlinear, positive monotonic, and positively accelerated with the RRCS2 the steepest once again.

For both standards, the transformed verbal scalings and the RRCS are in closer mutual agreement than they are with ME scales, but it must be noted that for the 1.55 standard they are closer to ME1 than ME2 is to ME1, and over the upper stimulus range for Standard 2.75 they are comparably close or closer to ME1 than is ME2. To summarize, if these alternative scaling procedures are used instead of ME1, they will yield comparable psychological magnitudes under some combinations of standard and scale steps that do not introduce too much asymmetry. They may be regarded as behaving like 'super-anchored' ME scales in that they differ from the unanchored ME2 even more than does the anchored ME1.

A magnitude estimation procedure gives a linear plot of log stimulus concentrations against log magnitudes, with, if necessary, a correction by the introduction of a threshold constant. It is found here that if the log magnitudes from the RRCS are used then the plots are linear; on this criterion RRCS are a sort of ME.

(3) The RRCS and the verbal scale can all be treated as equal interval category scales; if this is done then any information contained in the numerical labeling of the RRCS is lost. Hence, if, in fact, these numberings do not affect category usage the verbal and RRCS scales will be linear on one another when treated as equal interval, and both should exhibit the same sort of curvilinearity when plotted against magnitude estimations. The results are summarized in Table 5, and it is found on plotting that

the value of the standard strongly dominates any labeling effects when log stimulus intensity is taken as abscissa and the category steps as ordinate. The relations are positive monotonic with a slight positive acceleration, the slopes being steeper for the standard of 1.55. To a first approximation the categories, whether verbally or numerically labeled, correspond to equal log-stimulus differences over the lower part of the intensity range; the picture is not as clear-cut for more intense stimuli where the verbal scales apparently show more susceptibility to the differential effects of the two standards.

(4) Magnitude estimation procedures are usually adopted as a means towards estimating the exponent of the psychophysical power function. As we have shown that RRCSs and transformed verbal scales can approximate to ME, it remains to see what values of the exponent, *b*, are obtained from the alternative procedures, and the range of intersubjective differences shown.

Table 6 shows the values of *b* based on pooled data and the range of individual *b* values for all the 10 conditions examined.

It is seen that the P.S.E. values are higher for the RRCS than for the ME procedures; in other words the constant error is much less, and in this respect the RRCS procedure is to be preferred. The effect of the value of the standard upon the exponent is quite marked for the ME methods, but is small for the RRCSs where the asymmetry dominates. Under all conditions the interindividual differences in the *b* values are considerable, and span a similar range, with perhaps more spread for the ME. The practice of quoting a single exponent value for a sensory

Table 5
Mean Category Values, Treating all Scales as Equal Interval with Steps 1 through 11, for the Relative Perceived Stimulus Intensities for Each Scale Type and Standard Stimulus Concentration. All Flowmeter Settings as given in Table 1 are used, including Setting 1 which is Unreliable, for Illustration.

Flowmeter setting stimulus conc	Verbal Scale		RRCS1		RRCS2	
	standard stimulus at		standard stimulus at		standard stimulus at	
	1.55	2.75	1.55	2.75	1.55	2.75
1	3.04	2.86	2.98	2.76	2.76	2.54
2	3.45	2.98	3.06	2.76	2.92	2.98
3	4.33	3.76	4.35	3.56	4.32	3.64
4 ^w	5.42	4.92	5.70	4.88	5.28	4.62
5	7.23	5.31	6.84	5.24	6.58	5.35
6 ^x	6.94	5.52	7.02	6.20	7.04	6.02
7	7.74	5.64	7.44	6.38	7.39	6.22
8	8.21	6.67	8.14	7.12	7.66	7.52
9	8.92	7.18	8.20	7.70	8.48	8.06

^w This setting corresponds to standard 1.55×10^{-1} mg/liter.

^x This setting corresponds to standard 2.75×10^{-1} mg/liter.

Table 6
Parameter Values Associated with Each Combination of a Scale Type and Concentration of Standard Odor.

Scale Type	Std	b	P.S.E.	Range of b Values
Verbal	1.55	.554*	8.67	—
Verbal	2.75	.508*	7.83	—
RRCS1	1.55	.622	9.18	.29 to 1.01
RRCS1	2.75	.600	10.12	.20 to 1.16
RRCS2	1.55	.834	9.69	.33 to 1.46
RRCS2	2.75	.804	11.68	.20 to 1.20
ME1	1.55	.696	6.98	.22 to 1.44
ME1	2.75	.963	9.69	.31 to 1.81
ME2	1.55	.684	4.94	-.32 to 1.71
ME2	2.75	.844	6.08	.39 to 1.49

* These values are the consequence of putting $b = +1.0$ in the transformation by equal ratio steps. All exponents have been obtained by a least squares linear plot of $\log \psi / \log \phi$.

continuum must therefore be regarded with some misgiving. The values for b for olfaction, using other substances than Eugenol, which have been published previously, are lower than the range .6-8 obtained here; this may be due to scaling methods, to stimulus characteristics, and to our method of presentation, which is better controlled than in other studies.

It is seen that RRCS can be used for exponent determination, and that the resultant values will be similar, considering the large intersubjective variance of this parameter.

CONCLUSIONS

Some of the empirical characteristics of RRCSs have been examined on a sensory continuum that is taken to be prothetic.

When an RRCS is treated as a category scale, it behaves like a verbal category scale up to a linear transformation. When it is treated like an abbreviated form of magnitude estimation it behaves like ME. These differences in treatment are entirely scoring conventions employed by the E; for the S using the RRCS only one response strategy is available at one time. In this respect, therefore, the results are consistent with Torgerson's (1960) view that Ss can only make one sort of judgment, and the reported differences between scale types are the consequence of transformations employed by E in data processing.

The RRCSs used here were so constructed that over at least half their range of category steps they were consistent in numerical labeling with an equal ratio scale. This, taken with the particular stimulus concentration series used, facilitates the use of an RRCS in a way that is similar both to the Parducci transformation of the verbal scale and to the ME procedures. Empirically we have constructed cases here that, within the limits of measurement error, can be regarded as intermediate or bridging cases between category scaling and ME procedures. Obviously the

fact that this can be done at all indicates the need for some more extensive examination of RRCSs with a variety of numerical range assignments to steps. We venture to conclude that the nature of the relationships between category scaling and ME is not a closed and settled matter.

REFERENCES

- EISLER, H. Empirical test of a model relating magnitude and category scales. *Scandinavian Journal of Psychology*, 1962a, 3, 88-96.
- EISLER, H. On the problem of category scales in psychophysics. *Scandinavian Journal of Psychology*, 1962b, 3, 81-87.
- EISLER, H. Magnitude scales, category scales, and Fechnerian integration. *Psychological Review*, 1963, 70, 243-253.
- ENGEN, T., & PFAFFMAN, C. Absolute judgements of odor intensity. *Journal of Experimental Psychology*, 1959, 58, 23-26.
- ENGEN, T., & PFAFFMAN, C. Absolute judgements of odor quality. *Journal of Experimental Psychology*, 1960, 59, 214-219.
- GALANTER, E. H., & MESSICK, S. The relation between category and magnitude scales of loudness. *Psychological Review*, 1961, 68, 363-372.
- HARPER, R., & STEVENS, S. S. Subjective hardness of compliant materials. *Quarterly Journal of Experimental Psychology*, 1964, 26, 204-215.
- HELM, C. E., MESSICK, S., & TUCKER, L. R. Psychological models for relating discrimination and magnitude estimation scales. *Psychological Review*, 1961, 68, 167-177.
- JONES, F. N. Scales of subjective intensity for odors of diverse chemical nature. *American Journal of Psychology*, 1958a, 71, 305-310.
- JONES, F. N. Subjective scales of intensity for three odors. *American Journal of Psychology*, 1958b, 71, 423-425.
- OUGH, C. S., & STONE, H. An olfactometer for rapid and critical odor measurement. *Journal of Food Science*, 1961, 26, 452-456.
- PARDUCCI, A. Range-frequency compromise in judgement. *Psychological Monographs*, 1963, 77 (2, Whole No. 565).
- PARDUCCI, A. Category judgment: A range-frequency model. *Psychological Review*, 1965, 72, 407-418.
- REESE, T. S., & STEVENS, S. S. Subjective intensity of coffee odor. *American Journal of Psychology*, 1960, 73, 424-428.
- SANDUSKY, A., & PARDUCCI, A. Pleasantness of odor as a function of the immediate stimulus context. *Psychonomic Science*, 1965, 3, 231-232.
- STEVENS, S. S., & GALANTER, E. H. Ratio scales and category scales for a dozen perpetual continua. *Journal of Experimental Psychology*, 1957, 54, 377-411.
- SVENSON, O. On the estimation of darkness by direct scaling methods. *Reports of Psychological Laboratory of the University of Stockholm*, 1967, No. 234.
- TORGERSON, W. S. Quantitative judgement scales. In H. Gulliksen and S. Messick (Eds.), *Psychological scaling: Theory and applications*. New York: Wiley, 1960.

NOTES

1. The work reported here was supported by funds from the University Grants Committee of New Zealand, in grants made to R. A. M. Gregson and to M. J. Mitchell.

2. Address: Department of Psychology, University of Canterbury, Christchurch, New Zealand.

(Accepted for publication February 19, 1969.)