Perception & Psychophysics
1990, 48 (6), 551-558

Pervasiveness and magnitude of context effects: Evidence for the relativity of absolute magnitude estimation

HUGH J. FOLEY
Union College, Schenectady, New York

DAVID V. CROSS
State University of New York, Stony Brook, New York

and

JENNIFER A. O'REILLY
State University of New York, Albany, New York

To test the assertion that absolute magnitude estimation serves to minimize context effects, two experiments were conducted in which area stimuli were judged under differing conditions. In Experiment 1, four groups of subjects made magnitude estimations of triangles ranging in area from 1.5 to 3,072 cm^2. No standard or modulus was used, and instructions were similar to those used in absolute magnitude estimation experiments. Each group first judged a different subrange of the stimuli (1.5-24; 48-768; 6-96; or 192-3,072 cm^2) before making judgments of the remaining stimuli. In Experiment 2, two groups of subjects made magnitude estimations of triangles ranging in area from 1.5 to 12,288 cm^2, with each group first judging a different subrange of stimuli (1.5-24 cm^2 or 768-12,288 cm^2). The design and instructions were virtually identical to those used in absolute magnitude estimation experiments. Our results indicate that the wording of the instructions is not crucial and that judgments are influenced in two ways that are not predicted by proponents of absolute magnitude estimation. First, the power functions fit to the initially presented subranges (e.g., 1.5-24 cm^2), which were judged without contextual effects produced by previously presented stimuli, were inconsistent with one another. Second, judgments of the remaining stimuli were influenced by the subrange of stimuli judged initially. The prevalence of context effects in both experiments, in spite of instructional differences, suggests that although one should avoid using a standard and modulus, there is little else to be gained by adopting the absolute magnitude estimation procedure.

Gescheider (1988) has characterized two different approaches to contextual effects in psychophysical judgments. One approach, which is more cognitive (see Mellers, 1983a, 1983b; Mellers & Birnbaum, 1982), views contextual effects as omnipresent and unavoidable, in fact as integral and interesting components of judgment. The other approach, which is more sensory (see Zwislocki, 1983a; Zwislocki & Goodman, 1980), views contextual effects as a nuisance that may be avoided. Regardless of the orientation they favor, researchers in psychophysics are often interested in minimizing influences of one sort or another, and one potential means of doing so may be the use of absolute magnitude estimation proposed by Zwislocki and Goodman (1980). Certainly a number of researchers have begun to use the method in the collection of their data (Collins & Gescheider, 1989; Hellman & Meiselman, 1988; Verillo, 1983; Zwislocki, 1983b). What remains to be seen is the extent to which this approach has the effects desired by its originators.

Using instructions with an emphasis on independent judgment of each sequentially presented stimulus, and a procedure characterized by the lack of a standard or modulus, Zwislocki and Goodman (1980) believe that one can arrive at an absolute scale. An absolute scale "implies a fixed unit and, therefore, an absolute coupling between numerals and psychological magnitudes" (Zwislocki & Goodman, 1980, p. 28). The absolute scale is more restrictive than Stevens's ratio scale (Stevens 1946, 1951,

The authors wish to acknowledge the careful reading of earlier versions of this manuscript by David Emmerich, Mary Ann Foley, George Gescheider, Rhona Hellman, Neil Macmillan, Barbara A. Mellers, Mimi Teghtsoonian, and three anonymous reviewers. Portions of the research were supported by a Faculty Research Grant from Union College. Some of these results were reported at the 1989 Eastern Psychological Association conference. Requests for information about these studies should be directed to Hugh J. Foley, Department of Psychology, Union College, Schenectady, NY 12308. Requests for data and information can also be directed to FOLEYH@UNION.

1959), because not only is a true zero known, but the "units of measurement cannot be designated arbitrarily by the experimenter, so that even a multiplicative transformation is excluded" (Zwislocki, 1983a).

Zwislocki and Goodman (1980) suggest five propositions for testing the "*existence* and robustness of absolute scales" (p. 30; emphasis added), but they never test their model formally. It is somewhat surprising that none of the exponents from their experiment are reported, making it difficult to determine their stability, and Zwislocki and Goodman do not report a single instance of hypothesis testing. That is, although they readily state criteria by which one might determine that an absolute scale has been attained, they do not suggest any but the most casual means of testing whether the criteria have been met. To provide a more rigorous test of the assertions made by Zwislocki and Goodman, we will make extensive use of typical hypothesis testing procedures.

We will focus our experiments on Zwislocki and Goodman's (1980) fourth proposition, namely that "the scale does not depend on the intensity of the first stimulus presented in absolute magnitude estimation" (p. 30). To investigate this proposition, we varied the order in which particular subsets of a range of stimuli were presented to subjects. In the first experiment, groups of subjects first made absolute magnitude estimations of five stimuli from one of four subranges (1.5-24, 48-768, 6-96, or 192-3,072 cm^2) before judging all of the remaining stimuli from the other three subranges. In the second experiment, groups of subjects first made absolute magnitude estimations of five stimuli from one of two subranges (1.5-24, or 768-12,288 cm^2) before judging the remaining stimuli, which ranged from 1.5-12,288 cm^2.

To the extent that Zwislocki and Goodman (1980) are correct about the nature of magnitude estimation, when we employ their procedures, we would expect constant slopes and intercepts in the simple linear model:

$$\log R = \beta \log S + \log k.$$

We would also expect that judgments of a particular area stimulus would be the same regardless of the subrange first presented. If a relativistic approach is correct, then we would expect to find the judgments of identical stimuli to be more malleable, leading to different slopes, intercepts, or both.

EXPERIMENT 1

Method

Subjects. There were 40 undergraduates who participated in this experiment for course credit. All of the subjects were naive with respect to psychophysical experiments.

Materials. The stimuli were white triangles on 35-mm slides, presented on the dark background of a square projection screen which measured 1.22 m per side. There were 10 unique triangles, with each triangle twice as large as the preceding triangle, for a total range of 512:1. The stimuli were projected onto the screen from two different positions; from the close position, the stimuli ranged from 1.5 to 768 cm^2, and from the far position the stimuli ranged from 6 to 3,072 cm^2. It should be noted that there was, in fact, substantial overlap between stimuli at the two projector positions. The stimuli common to both positions were 6, 12, 24, 48, 96, 192, 384, and 768 cm^2. So the only unique stimuli were 1.5 and 3 cm^2 for the close position, and 1,536 and 3,072 cm^2 for the far position.

Procedure. The subjects were first presented with a random order of 10 lines that varied in length from 5 to 110 cm in a geometric progression. Instructions were chosen that would closely replicate those used by Zwislocki and Goodman (1980). The subjects were instructed to give a number to match the length of the line, and were told that they could use any numbers they wanted, but that zero and negative numbers would not be needed (because a stimulus of some length would always be present). The subjects were expressly told that fractions or decimals were permissible, and that they should simply assign numbers that they felt were appropriate for each stimulus presented. The subjects were told that they should not feel constrained to use a common metric such as inches, centimeters, or any other common metric. On completion of the line-length judgment task, the subjects were asked if they had any questions about the judgment process, prior to being instructed about the area judgment task.

The subjects were informed that they would be using numbers to judge the area of several triangles, and that they should use numbers in just the way that they had used them in judging the lengths of the lines, except that they would be judging areas.

To assess the effects of the initial stimuli presented, half the subjects first saw the stimuli presented at the close projector position; then the projector was moved, and they judged the stimuli from the far projector position. The remaining subjects first saw the stimuli at the far position, and then at the close projector position. Because we were concerned that the shift in projector position might signal to the subject that the nature of the stimuli had changed, the stimuli were presented randomly within two blocks at each projector position. Half of the subjects first judged a block of the smaller 5 of the 10 areas (presented twice each), followed by the larger 5 of the 10 areas. The order of the blocks was reversed for the remaining subjects. The first stimulus presented from any block was chosen from the middle of the range for that block.

So subjects were presented with 20 stimuli, after which the projector was moved and the remaining stimuli presented. With the counterbalancing of projector position, there were four groups in the experiment. Group 1 judged the area stimuli in the order 6-96, 192-3,072, 48-768, and 1.5-24 cm^2. Group 2 judged the stimuli in the order 1.5-24, 48-768, 192-3,072, and 6-96 cm^2. Group 3 judged the stimuli in the order 192-3,072, 6-96, 48-768, and 1.5-24 cm^2. Group 4 judged the stimuli in the order 48-768, 1.5-24, 192-3,072, and 6-96 cm^2.

Results and Discussion

Analyses of log responses showed no differences between judgments of the first and second presentations of the stimuli, so all analyses are based on the average of the log responses to the two stimuli at each level of area.

Because of the fact that the stimuli at each projector position were presented in two blocks, with a different set of stimuli in each block, it was possible to examine the effects of context for the two sets of stimuli within subjects, but without the complication of a shift in projector position. Using only the responses to stimuli presented prior to moving the projector, the effects of context were evident. Figure 1 shows the responses to the smaller and larger stimulus sets at the close projector position for Groups 2 and 4. Group 2 saw the smaller set first, then the larger set, and Group 4 saw the two sets in the reverse order. Figure 2 shows responses to the two sets of stimuli

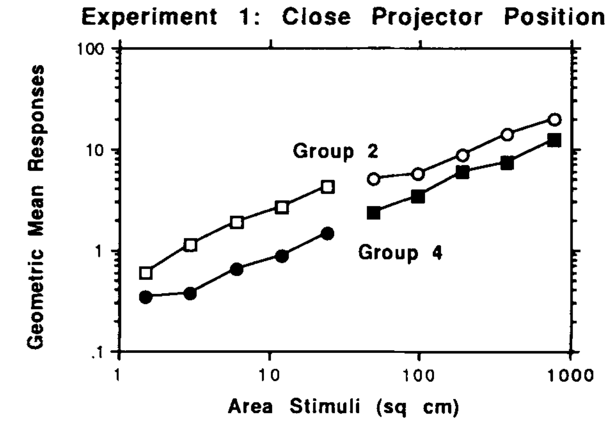

Figure 1. Geometric mean responses for the two groups that initially judged stimuli at the close projector position. Responses are shown as a function of the area stimuli in log-log coordinates. The square symbols indicate the stimuli judged first for each group, and the circle symbols indicate the second block of stimuli judged. Group 2 (open symbols) judged the smaller block of stimuli first (1.5-24 cm²), and Group 4 (filled symbols) judged the larger block of stimuli first (48-768 cm²).

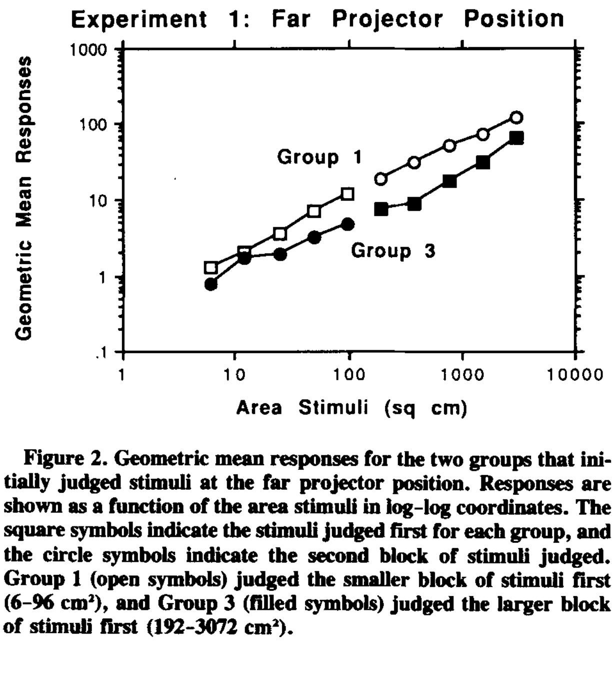

Figure 2. Geometric mean responses for the two groups that initially judged stimuli at the far projector position. Responses are shown as a function of the area stimuli in log-log coordinates. The square symbols indicate the stimuli judged first for each group, and the circle symbols indicate the second block of stimuli judged. Group 1 (open symbols) judged the smaller block of stimuli first (6-96 cm²), and Group 3 (filled symbols) judged the larger block of stimuli first (192-3072 cm²).

presented at the far projector position. Group 1 saw the smaller set first, and Group 3 saw the larger set first.

The relativity of the subjects' judgments becomes apparent in two ways. First, there is a noticeable discontinuity between judgments of the smaller set (by Groups 1 and 2, open squares) and the larger set (by Groups 3 and 4, filled squares), when the subjects had not yet been exposed to other areas. If magnitude estimations are absolute, then one would expect that these judgments of initially presented blocks of stimuli, relatively uncontaminated by contextual effects, should fall along the same line of best fit. Obviously, they do not. Second, identical stimuli are judged quite differently when judged subsequent to observing the other stimulus set, which shows a strong influence of the first stimulus set judged. A comparison of the responses to the second stimulus set (circles) in both figures with the responses to identical areas presented first (squares) shows strong contextual effects on judgments.

For the close projector position (Figure 1), the subjects who saw the smaller set of stimuli first (Group 2) gave significantly larger responses overall ($M = .800$) than did the subjects who saw the larger set of stimuli first (Group 4, $M = .269$) [$F(1,18) = 5.1$]. As can be seen in Table 1, these differences are apparent in the slopes and intercepts of functions fit to the five stimuli of the smaller and larger stimulus sets. The intercepts for Group 2 approach being significantly larger than the intercepts for Group 4 [$F(1,18) = 3.18, p < .09$], which is consistent with the results for log responses. For the exponents, there is an interaction between stimulus set and the order in which the sets were presented [$F(1,18) = 4.52$]. This effect might be due to the trend for exponents to be larger earlier in the experiment.

At the far projector position (Figure 2), it was again the case that the subjects exposed first to the smaller stimulus set (Group 1) produced larger judgments ($M = 1.131$) than did the subjects who judged the larger set first (Group 3, $M = .601$) [$F(1,18) = 7.89$]. As was the case for the close projector position, the subjects tended to produce a larger exponent for the stimuli presented earlier in the sequence, producing an interaction between responses to the smaller and larger set and the order in which the sets were presented [$F(1,18) = 4.08$]. For the larger stimulus set, the intercepts for the group getting the smaller stimulus set first are larger than for the group getting the larger set first, but this is not true for the smaller stimulus set, producing another interaction [$F(1,18) = 5.03$].

In a comparison of both projector positions, the exponents for the far position were significantly larger ($M = .72$) than those for the close position ($M = .58$) [$F(1,36) = 10.95$], which is consistent with the results of earlier experimentation (Foley, Cross, Foley, &

Table 1
Slopes and Intercepts for Both the Smaller and Larger Stimulus Sets, Including Only Responses Prior to Moving the Projector: Experiment 1

	Smaller Set		Larger Set	
	Slope	Intercept	Slope	Intercept
	Close Projector Position			
Group 2 (SM-LG)	.683	−.296	.531	−.246
Group 4 (LG-SM)	.531	−.605	.588	−.613
	Far Projector Position			
Group 1 (SM-LG)	.814	−.542	.662	−.227
Group 3 (LG-SM)	.615	−.526	.808	−1.056

Note—SM and LG refer to the order in which subjects saw the smaller and larger stimulus set, respectively.

Reeder, 1983). The interactions found at both projector positions combined to produce an interaction between stimulus set size and order of presentation [$F(1,36) = 7.91$], caused by the tendency of subjects to produce steeper slopes for the first stimulus set presented. At both positions, subjects who saw the smaller set first produced larger responses, which led to significantly larger intercepts ($M = -.327$) than those found for subjects who saw the larger set first ($M = -.700$) [$F(1,36) = 4.59$]. For the intercepts, there was an interaction consistent with that found for the exponents, with subjects producing a smaller intercept for the stimuli they saw first [$F(1,36) = 4.80$].

Subjects do tend to give larger responses to larger stimuli, so that the smaller stimulus set received responses that were significantly lower than responses to the larger stimulus set, and responses to stimuli presented at the close projector position were smaller than responses to stimuli presented at the far projector position. However, there is a considerable overlap in judgment for the larger stimuli in the smaller set and the smaller stimuli in the larger set, which is apparent at both projector positions. It is important to note that this was the case even when subjects had not yet been exposed to any other stimuli. It was also the case that stimuli common to the two projector positions were judged quite differently. So whether we restrict ourselves to stimuli judged before a shift in projector position, or look at all judgments of stimuli common to both projector positions, the picture that emerges is one of relativity of judgment.

EXPERIMENT 2

Reviewers of an earlier version of this paper felt that we had not replicated the absolute magnitude instructions of Zwislocki and Goodman (1980), and they were also uncomfortable with the shift in projector position (in spite of the preceding analyses, which were done solely *within* a particular projector position, prior to a shift in position). Experiment 2 was conducted to address some of their concerns.

To the best of our knowledge, there is no experimental evidence that the wording of the instructions is crucial. By that we mean that there has been no study in which a systematic manipulation of absolute magnitude estimation instructions has produced an accompanying change in results. In fact, as Zwislocki and Goodman (1980) point out, Ward (1973) obtained results consistent with an absolute scale even though he used a standard and a modulus. Nonetheless, two reviewers felt that it was crucial to tell subjects to match their *impression* of number size with their *impression* of the stimulus. We have done so in Experiment 2, so this pair of experiments represents a test of the importance of the wording of instructions. To the extent that we obtain results similar to those obtained in Experiment 1, our data would argue that the slight instructional differences between our experiments and other studies of absolute magnitude estimation are inconsequential.

Method

Subjects. There were 20 undergraduates who participated in this experiment for $3. All of the subjects were naive with respect to psychophysical experiments.

Materials and Procedure. The stimuli were identical to those used in Experiment 1, but a larger square projection screen was used (2.44 m per side), and the far projector position was changed so that the range of stimuli presented at that position was from 24 to 12,288 cm^2. Stimuli at the close projector position again ranged from 1.5 to 768 cm^2.

To avoid moving the projector, two projectors were used. The subjects were seated near the close projector, and the experimenter controlled the stimulus presentation from one location with remote advance switches so that there were no cues to changes between projector positions.

There were two groups of subjects, determined by the order of the particular block of stimuli judged first. The subjects in Group 1 saw the 5 stimuli from the smallest possible block of stimuli first (1.5–24 cm^2), followed by the 5 stimuli from the largest possible block (768–12,288 cm^2). The subjects in Group 2 saw the two blocks in reverse order. After these first 10 stimuli, both groups saw stimuli presented in identical orders for the remainder of the experiment. The subjects next saw the following blocks of areas: 48–768, 24–384, 768–12,288, 1.5–24, and 48–768 cm^2. Within each block, the 5 stimuli were presented in a randomly determined order, with the restriction that the first stimulus in a block was neither the largest nor the smallest stimulus within that block. Thus, the subjects made a total of 35 area judgments, and all subjects in both groups judged the same 35 stimuli.

As in Experiment 1, the subjects made line-length judgments first. They were instructed as follows:

> You will be shown a series of line lengths and asked to judge each of them with a number. Your task is to assign numbers to the line lengths such that your subjective impression of the magnitude of the number matches your subjective impression of the magnitude of the line length. You should concentrate on each line as it is presented, and not be concerned with the numbers you assigned to previous line lengths. You can use any positive numbers that seem appropriate to you—including whole numbers, fractions, and decimals. You should not think of any rules you might have learned about numbers, but should make your responses as spontaneously and quickly as possible. You should also not be at all concerned to make your judgments along dimensions with which you might be familiar (inches, centimeters, etc.).

After judging the 10 line lengths, subjects never had any questions, so they were instructed for the area stimuli as follows:

> As was the case for the line lengths, you should assign numbers to each of the triangle areas presented. You should again match your subjective impression of the magnitude of the number to your subjective impression of the magnitude of the area stimulus. Don't be concerned with the numbers you assigned to the line lengths, or to numbers assigned to previous area stimuli. You should simply concentrate on each area as it is presented to you. As before, you can use any positive numbers that seem appropriate to you. Your responses should again be quick and spontaneous, and you should not be at all concerned about making your judgments in familiar units (square centimeters, square inches, etc.).

Results and Discussion

Given the results of Experiment 1, we expected that the first block of stimuli the subjects judged would have an impact on their subsequent judgments. Thus, we expected to find a significant difference between the two groups when comparing their responses to identical stimuli. As was done in Experiment 1, all analyses were performed on logarithmically transformed data.

Figure 3 shows the responses to the first two blocks of stimuli presented to the two groups. As was the case in Experiment 1, Group 1 (which first judged the smaller stimuli and then the larger stimuli) tended to give larger judgments overall ($M = .90$) relative to Group 2 ($M = .48$) [$F(1,18) = 5.32$]. These results are certainly consistent with the notion that for the first stimuli they judge, subjects choose magnitude estimation responses that tend to be toward the middle of their response range. In other words, they seem to have a somewhat fixed range of responses in mind, and they try to leave themselves room to respond to both smaller and larger stimuli than those initially presented. Even though instructed to respond to each stimulus independently, subjects seek to be consistent across their responses, which produces an overall difference between groups that start with different ranges of stimuli.

The differences in area judgments made by the two groups were not consistent across both blocks, producing an interaction [$F(9,162) = 2.92$]. As subsequent analyses indicated, the differences between the two groups were more pronounced with the larger stimuli. Even though Group 1 produced larger judgments for the five smallest stimuli ($M = .07$) than did Group 2 ($M = -.17$), this difference was not significant [$F(1,18) = 1.72, p < .21$]. For the five largest stimuli, Group 1 did produce significantly larger judgments ($M = 1.73$) than did Group 2 ($M = 1.14$) [$F(1,18) = 7.44$].

These effects are apparent even when one considers judgments of the first presentations of all 10 stimuli at each projector position. Again, Group 1 produced significantly larger judgments ($M = .86$) than Group 2 ($M = .50$) [$F(1,18) = 4.28$], although the 20 stimuli were identical. These differences were again less pronounced for the smaller stimuli ($M = .51$ and $.22$ for Groups 1 and 2, respectively) [$F(1,18) = 3.13, p < .09$], and more pronounced for the larger stimuli ($M = 1.21$ and $.78$, respectively) [$F(1,18) = 4.75$].

Table 2
Slopes and Intercepts from Experiment 2

	Smaller Set		Larger Set	
	Slope	Intercept	Slope	Intercept
	Close Projector Position			
Group 1 (SM-LG)	.670	−.452	.443	−.059
Group 2 (LG-SM)	.660	−.685	.374	−.241
	Far Projector Position			
Group 1 (SM-LG)	.469	−.227	.477	.070
Group 2 (LG-SM)	.536	−.649	.570	−.849

Note—SM and LG refer to the order in which subjects saw the smaller and larger stimulus sets, respectively.

Table 2 shows the slopes and intercepts for the subsets of stimuli judged by the two groups. The pattern of results is similar to that seen in Experiment 1, with generally larger intercepts for Group 1 ($M = -.167$) than for Group 2 ($M = -.606$) [$F(1,18) = 3.98, p < .06$]. The intercept differences parallel those obtained in the first experiment, and are consistent with the larger magnitude of responses found in Group 1. The magnitude of the slopes at the far projector position is smaller in Experiment 2 than in Experiment 1, possibly because of the increase in size of the stimuli at the far projector position in Experiment 2.

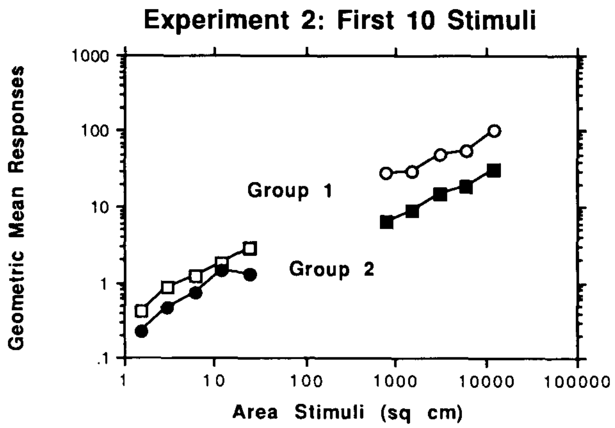

Figure 3. Geometric mean responses for the first 10 stimuli presented to the two groups. Responses are shown as a function of the area stimuli in log-log coordinates. The square symbols indicate the stimuli judged first for each group, and the circle symbols indicate the second block of stimuli judged. Group 1 (open symbols) judged the smaller block of stimuli first (1.5-24 cm²) and Group 2 (filled symbols) judged the larger block of stimuli first (768-12,288 cm²).

GENERAL DISCUSSION

Taken together, our two experiments show that the specific instructions and procedures espoused by Zwislocki and Goodman (1980) do not serve to minimize contextual effects, as Gescheider (1988) claims. Such instructional manipulations are often ineffective, because telling subjects to do something is no guarantee that they will comply. For instance, telling subjects to make their judgments relative to a standard and modulus does not mean that they will do so (Baird, Kreindler, & Jones, 1971; Hellman & Zwislocki, 1961; Ward, 1987). Telling subjects to judge each stimulus independently does not eliminate sequential effects (Ward, 1987). Finally, as we have shown, telling subjects to match their subjective impression of the magnitude of a stimulus with their subjective impression of the magnitude of a number does not lead to different responses than simply telling subjects to give a number that seems appropriate for each stimulus.

It appears that the absolute magnitude estimation procedure offers researchers little gain beyond the notion that one should avoid using a standard and a modulus, which has been known since the work of Hellman and Zwislocki (1961). Our results, and the results of others (Ward, 1987;

Zwislocki & Goodman, 1980) argue that the procedure does *not* minimize contextual effects.

What, then, are the differences between our studies and those of Zwislocki and Goodman (1980)? There are a number of differences between the two sets of studies, including the stimulus modalities used, the subject population, and slight instructional differences, among others. Ward (1987) and Marks (1988; Marks, Szczesuil, & Ohlott, 1986) have recently shown substantial contextual effects in judgments of auditory stimuli, suggesting that our results are probably not limited to judgments of area stimuli. However, it is more instructive to reexamine the results of Zwislocki and Goodman (1980) before deciding to attribute our contrasting conclusions to any of the differences above.

A Reevaluation of Zwislocki and Goodman (1980)

Two of Zwislocki and Goodman's (1980) five propositions for absolute scales revolve around the identity of production and estimation scales. An examination of their Figure 4 reveals that the functions were quite different, with the production slope about twice as steep as the estimation slope. This is true even though they chose a set of numerals as stimuli in the production task to match those given by subjects in the estimation task. In other words, they did everything possible to ensure the similarity of the production and estimation functions, and still they were not the same.

A more reasonable test of the absoluteness hypothesis might have been to use numerals larger than 10 for the production task. If the coupling between numerals and sensations is absolute, then the data of Figure 4 from Zwislocki and Goodman (1980) indicate that all of the responses should be greater than 65 dB (using the production function) or 80 dB (using the estimation function). We doubt seriously that this would have been the case. We would suggest that subjects asked to match numerals ranging from 10 to 1,000 would be more likely to match 1,000 with a 90- or 100-dB tone, and then match 10 with a much softer tone (e.g., 40 or 50 dB). Surely Zwislocki and Goodman are not arguing that magnitude production experiments should only be conducted using stimuli between .1 and 10?

Zwislocki and Goodman (1980) also argue that there should be no effect of the first stimulus in the series, yet they find a "small, but systematic, effect on the loudness function" for magnitude estimation. Interpolations from their Figure 6 reveal that this "small" difference can amount to a difference in mean judgment from about 17 to about 80 for an 80-dB sound. Effects of the first stimulus are also apparent in slopes fit to interpolations of their data, where groups getting 6, 36, and 78 dB first produced slopes of .53, .35, and .43, respectively. Although this effect is attenuated somewhat for magnitude production (their Figure 7), similar slope differences remain (about 1.0 and .72 for the groups getting .1 and 10 first, respectively).

A fourth proposition is that stimulus range does not affect the magnitude estimation function. The current experiments, as well as earlier research (Foley et al., 1983), demonstrate that not only the stimulus range, but also the positions of the stimuli in the total available range, appear to affect psychophysical judgments. What are the effects of range manipulations in the Zwislocki and Goodman (1980) study?

One group (L-H) was given a three-run session of a low range of stimuli (L, 6–54 dB), followed by four three-run sessions of a high range of stimuli (H, 30–78 dB), followed by three three-run sessions of low stimuli. The other group (H-L) had the opposite order of the H and L sequences. In arguing for the absoluteness of magnitude estimation, Zwislocki and Goodman (1980) attempt to demonstrate that range has no effect, and that shifting from one sequence to another has no effect. Shifting sequences ought most likely to be apparent in judgments of the *first* run of a new sequence, before the subject has had a chance to adapt to the new stimulus levels, but Zwislocki and Goodman only look at the last two runs of a shift session, or else average all three runs together. Nonetheless, Table 3 shows what appear to be striking and systematic range and adaptation effects in their data.

For instance, within each group, judgments of the low series produce steeper slopes than judgments of the high series do. This is true even across the two groups before they have been exposed to the other series (i.e., Session 1). In Session 2, where the first shift occurs, we see a change in function for both groups, but the H-L group did not shift fully to the slope an unadapted group might produce (e.g., the L-H group function in Session 1). Instead, they shift to an intermediate slope, which suggests some remaining influence of the initial session. Over time (comparing Sessions 2 and 5) the slopes for both groups shift toward their own initial slopes in Session 1, rather than toward the slopes of the opposite group in Session 1. Finally, in Session 6, there is a shift for both groups back to a slope near that found in Session 1. The malleability of the slopes is apparent throughout the table, even though the averaging process used by Zwislocki and Goodman

Table 3
Slopes Calculated from Data Interpolated from Zwislocki & Goodman (1980), Figures 8-11

		L-H Group		H-L Group	
Session	Run	Stimulus Type	Slope	Stimulus Type	Slope
1	1	L	.70	H	.42
1	2 & 3	L	.79	H	.55
2	2 & 3	H	.55	L	.60
5	all	H	.63	L	.52
6	all	L	.80	H	.51

Note—H and L refer to high and low decibel ranges, respectively.

(1980) serves to minimize the influence of shifts in stimulus range.

Can Psychophysicists Achieve an Absolute Scale?

The scaling typology specified by Stevens is not without critics, and the determination of the type of scale being used in a particular instance is problematic (Prytulak, 1975). If by an absolute scale, adherents to this approach mean only that subjects like to use a particular range of numbers, and that they will drift to that range regardless of instructions, then we have no real criticism of their position. However, Collins and Gescheider (1989) argue that their cross-modality matching results are not produced by subjects' simply using the same range of numbers to match both lines and tones, despite the fact that more than 78% of their subjects produced data that are totally consistent with such an interpretation.

Subjects are good at assigning numbers to stimuli such that smaller stimuli are assigned smaller numbers and larger stimuli are assigned larger numbers. As Parker, Casey, Ziriax, and Silberberg (1988) note, however, if subjects do nothing more than that, and stimuli are spaced systematically, correlation coefficients and coefficients of determination will be high. Their research shows that it is risky to employ such statistics to test the goodness of fit of a function, as many researchers do (e.g., Collins & Gescheider, 1989).

Our results are consistent with the notion that subjects have a common extraexperimental experience with magnitudes of numbers and experiences in several modalities (length, area, sound pressure, brightness, etc.), as Zwislocki and Goodman (1980) suggest. Subjects do generally use smaller numbers to judge smaller stimuli, and larger numbers to judge larger stimuli. Such common experience seems crucial to explaining the results of most psychophysical experiments, including the one in which subjects are called on to make only one judgment of a single stimulus (Stevens, 1975). On the other hand, within an experiment in which several stimuli are presented sequentially for judgment, each judgment is more likely an amalgam of the extraexperimental experience that the subject brings to the laboratory and also the intraexperimental influences that develop in the course of the experiment (Helson, 1964; Ward, 1987). This joint effect of such intra- and extraexperimental influences was noted by Stevens (1956), who reported that a subject "seems to be trying to say, for example, that such-and-such variable is not only five times louder than the standard but it is also a very loud tone indeed" (p. 7).

The power law proposed by Stevens is a good first-order approximation of the relationship between stimuli and responses, but with only two parameters available, contextual effects have to show themselves in the exponent and/or the intercept of the power function. As we have shown, there is substantial variability in both exponents and intercepts, depending on the context in which stimuli are presented. We would argue that even when the precautions of Zwislocki and Goodman are observed, the intraexperimental context effects are such robust second-order effects as to effectively preclude an absolute coupling between a perception and the subject's response.

Overall, then, not only our own data, but a more detailed inspection of Zwislocki and Goodman's (1980) data reveals that absolute magnitude estimation judgments are not absolute—using the criteria that Zwislocki and Goodman propose. Simply looking at graphs of data is often an insensitive way of detecting very real differences in the data. To support a claim that these differences are not real would require of Zwislocki and Goodman a more specific statement of a means of evaluating their criteria. Zwislocki (1983a) argues that even physical laws hold only under specified conditions, and that

> to demonstrate the realizability of a scale, it is sufficient to find one set of conditions under which the scale can be proved to exist. The discovery of circumstances that preclude the achievement of a scale does not constitute a proof of its general lack of validity. (p. 593)

Although this is true up to a point, we do not feel that our experiments or those of others (Marks, 1988; Ward, 1987) represent a substantial departure from the conditions found in Zwislocki and Goodman (1980), while our conclusions are radically different. If one goal of science is predictability, then we feel that a more viable approach to modeling psychophysical judgment is to add sufficient parameters to the simple power law, to enable it to predict judgment under a wide range of conditions. As Stevens (1951) says,

> a law would not be worth much if it worked only in Europe or only at sea level or only for circular motions or only for objects larger than elephants. The invariance of the law under a wide assortment of conditions is its source of power. (p. 20)

Attempts to model the judgment process more accurately (see, e.g., Lockhead & King, 1983), or to establish the nature of contextual influences on judgment (see, e.g., Anderson, 1975; Mellers, 1983a) should ultimately prove more useful than attempts to minimize the relativity of judgment.

REFERENCES

Anderson, N. H. (1975). On the role of context effects in psychophysical judgment. *Psychological Review*, **82**, 462-482.

Baird, J. C., Kreindler, M., & Jones, K. (1971). Generation of multiple ratio scales with a fixed stimulus attribute. *Perception & Psychophysics*, **9**, 399-403.

Collins, A. A., & Gescheider, G. A. (1989). The measurement of loudness in individual children and adults by absolute magnitude estimation and cross-modality matching. *Journal of the Acoustical Society of America*, **85**, 2012-2021.

Foley, H. J., Cross, D. V., Foley, M. A. & Reeder, R. (1983). Stimulus range, number of categories, and the "virtual" exponent. *Perception & Psychophysics*, **34**, 505-512.

Gescheider, G. A. (1988). Psychophysical scaling. *Annual Review of Psychology*, **39**, 169-200.

Hellman, R. P., & Meiselman, C. H. (1988). Prediction of individual loudness exponents from cross-modality matching. *Journal of Speech & Hearing Research*, **31**, 605-615.

HELLMAN, R. P., & ZWISLOCKI, J. J. (1961). Monaural loudness function at 1000 cps and interaural summation. *Journal of the Acoustical Society of America*, **35**, 856-865.

HELSON, H. (1964). *Adaptation level theory: An experimental and systematic approach to behavior*. New York: Harper & Row.

LOCKHEAD, G. R., & KING, M. C. (1983). A memory model of sequential effects in scaling tasks. *Journal of Experimental Psychology: Human Perception & Performance*, **9**, 461-473.

MARKS, L. E. (1988). Magnitude estimation and sensory matching. *Perception & Psychophysics*, **43**, 511-525.

MARKS, L. E., SZCZESUIL, R., & OHLOTT, P. (1986). On the cross-modal perception of intensity. *Journal of Experimental Psychology: Human Perception & Performance*, **12**, 517-534.

MELLERS, B. A. (1983a). Evidence against "absolute" scaling. *Perception & Psychophysics*, **33**, 523-526.

MELLERS, B. A. (1983b). Reply to Zwislocki's views on "absolute" scaling. *Perception & Psychophysics*, **34**, 405-408.

MELLERS, B. A., & BIRNBAUM, M. H. (1982). Loci of contextual effects in judgment. *Journal of Experimental Psychology: Human Perception & Performance*, **8**, 582-601.

PARKER, S., CASEY, J., ZIRIAX, J. M., & SILBERBERG, A. (1988). Random monotone data fit simple algebraic models: Correlation is not confirmation. *Psychological Bulletin*, **104**, 417-423.

PRYTULAK, L. S. (1975). Critique of S. S. Stevens' theory of measurement scale classification. *Perceptual & Motor Skills*, **41**, 3-28.

STEVENS, S. S. (1946). On the theory of scales of measurement. *Science*, **103**, 677-680.

STEVENS, S. S. (1951). Mathematics, measurement, and psychophysics. In S. S. Stevens (Ed.), *Handbook of experimental psychology* (pp. 1-49). New York: Wiley.

STEVENS, S. S. (1956). The direct estimation of sensory magnitudes—loudness. *American Journal of Psychology*, **69**, 1-25.

STEVENS, S. S. (1959). Measurement, psychophysics and utility. In C. W. Churchman & P. Ratoosh (Eds.), *Measurement: Definitions and theories* (pp. 18-63). New York: Wiley.

STEVENS, S. S. (1975). *Psychophysics: Introduction to its perceptual, neural and social prospects*. New York: Wiley.

VERILLO, R. T. (1983). Stability of line-length estimates using the method of absolute magnitude estimation. *Perception & Psychophysics*, **33**, 261-265.

WARD, L. M. (1973). Repeated magnitude estimations with a variable standard: Sequential effects and other properties. *Perception & Psychophysics*, **13**, 193-200.

WARD, L. M. (1987). Remembrance of sounds past: Memory and psychophysical scaling. *Journal of Experimental Psychology: Human Perception & Performance*, **13**, 216-227.

ZWISLOCKI, J. J. (1983a). Absolute and other scales: Question of validity. *Perception & Psychophysics*, **33**, 593-594.

ZWISLOCKI, J. J. (1983b). Group and individual relations between sensation magnitudes and their numerical estimates. *Perception & Psychophysics*, **33**, 460-468.

ZWISLOCKI, J. J., & GOODMAN, D. A. (1980). Absolute scaling of sensory magnitudes: A validation. *Perception & Psychophysics*, **28**, 28-38.

NOTE

1. All hypothesis testing was conducted with $\alpha = .05$.

(Manuscript received June 28, 1989;
revision accepted for publication July 24, 1990.)