

Commentary on Jacobs and Michaels (2001): Calibration and perceptual learning in event perception

GEOFFREY P. BINGHAM
Indiana University, Bloomington, Indiana

DANIEL S. McCONNELL
*University of Pennsylvania,
Philadelphia, Pennsylvania*

and

MICHAEL M. MUCHISKY
Morningside College, Sioux City, Iowa

Jacobs and Michaels (2001) have argued that increased precision in judgments of the viewing distance to a perceived event should be attributed in part to perceptual learning. They found that observers used feedback to attune to the appropriate information variables gradually. McConnell, Muchisky, and Bingham (1998) had found that observers used feedback to calibrate event-specific scaling coefficients, that the calibration of one type of event generalized to other types, and that calibration occurred suddenly. We argue that Jacobs and Michaels must be partially correct and that, in our experiments, both calibration and perceptual attunement were required for accurate and precise judgments.

McConnell, Muchisky, and Bingham (1998) investigated whether gravitationally governed forms of motion and timing in events could be used to perceive absolute scale (either size or distance) in events. This had been examined in previous studies that had provided evidence in support of the hypothesis. However, only one type of event had been investigated in each study. This allowed the possibility that subjects simply detected a temporal property (e.g., an event duration, a peak velocity, or a mean velocity) that covaried with the spatial scale. To control for this, we asked observers to judge absolute spatial scale (either size or distance) in a variety of different gravitationally governed events (i.e., free fall, balls rolling down inclines at different angles, and pendula of different lengths). An event-specific scaling coefficient was required to map temporal properties to scaled spatial properties. The point was that a different scaling coefficient would be required for each different type of event, so that a simple covariation between spatial scale and temporal properties was broken. The idea was that the form of the

motion in the event would specify to an observer the appropriate scaling relation that would then map the available temporal property to the requisite spatial property.

Given the design of our study, there were a number of different temporal properties that could be used in this way. We performed analyses by assuming the use of event durations. For each event, we computed a physically correct coefficient that transformed times to either object sizes (Experiment 1) or viewing distances (Experiment 2). We derived physical coefficients by measuring times in the displays and regressing them on the corresponding sizes or distances. We did the same by using judged sizes or distances to derive the coefficient values represented by the observers' performance. We then compared the physical and the perceptual coefficients. It is possible that some or all of the observers used a different temporal property than event duration. They might equally well have used peak velocity or average velocity.¹ If observers used mean velocity instead of event duration, a different set of coefficients would be entailed, but those coefficients would covary perfectly with those for event duration, and the analysis would otherwise be the same. We made no attempt to discriminate among temporal properties that could be used by observers in this way. However, our multievent design did explicitly allow us to discriminate between the use of properties that specified absolute spatial scale and those that did not. For instance, if observers had simply assigned numbers to detected durations or to detected velocities, this would have been revealed by the resulting poor performance. That is, judgments would not have covaried with the actual spatial scale.

In our first experiment, observers judged the size of the ball in the various events. We obtained results that were comparable with those in previous studies. On average, judgments covaried with actual sizes, but performance was highly variable, owing to a large amount of between-subjects variability. In a second experiment, we tested the effect of feedback in calibrating the judgments and reducing between-subjects variability. We found one-trial learning in which the variability collapsed to yield relatively accurate and precise performance. Also, feedback provided for only one type of event was found to generalize across all of the other types of events. We computed both physical and perceptual event coefficients and found that they correlated very well. However, we also found that the ratios of the physical coefficients correlated with the ratios of perceptual coefficients with a slope (.96) closer to 1. Even with calibration, relative scale was perceived more accurately than absolute scale.

We had suggested that observers' judgments correctly reflected the relations among events (as reflected in relative scaling and in the ratios of the coefficients) but that feedback was required to calibrate the absolute values of the coefficients. Akin to the strategy used by Raibert

The writing of this commentary was supported by NEI Grant R01 EY11741-01A1, awarded to the first author. Correspondence concerning this article should be addressed to G. P. Bingham, Department of Psychology, 1101 East Tenth Street, Indiana University, Bloomington, IN 47405 (e-mail: gbingham@indiana.edu).

(1986) in the organization and design of stable and effective robots, we suggested that perception is attuned to approximate linear regularities without sensitivity to the effects of variations in friction or mass. Feedback can be used to achieve rapid, one-trial calibration that then takes into account these other factors. (See also Pizlo and Rosenfeld, 1992, and Pizlo, Rosenfeld, and Weiss, 1997, for an engineering approach to the derivation of “approximate invariants” that is similar in spirit.)

Jacobs and Michaels (2001) noted that we had failed to perform an analysis of the perceptual event-specific coefficients obtained from judgments made before feedback was provided. In this case, we would predict that the ratios among the coefficients would be the same as that found after feedback had been provided, because calibration is not required for relative scaling. In our second experiment, observers had judged displays of only one type of event before they received feedback, so this data would not be appropriate for the analysis. However, observers had judged multievent displays without feedback in Experiment 1. They had judged object size rather than viewing distance, so the values of the physical coefficients would be different from those in Experiment 2. Also, there were more events than in Experiment 2. Table 1 contains the physical and perceptual coefficients derived from the data in Experiment 1. These were derived in the same way as those shown in Table 5 of McConnell et al. (1998). The regression of physical coefficients on coefficients derived from judgments yielded an r^2 (.99) that was quite high, although the slope (.77) was low. The latter reflects the inaccuracy of the coefficients, whereas the former reflects the accuracy of the relations among them. The respective ratios are shown in Table 2. We include only ratios for events used in both Experiments 1 and 2, to keep Table 2 comparable to Table 6 in McConnell et al. The regression of physical ratios on ratios derived from judgments yielded a high r^2 (.99) once again as shown in Figure 1, but this time the slope (.91) was close to 1, as was predicted. Using the ratios in Table 6 of McConnell et al., we computed the mean absolute difference between the corresponding ratios (physical vs. those derived from judgments) in proportion to the mean physical value. We did the same using the data in Table 2. The results were 9% and 8.7%, respectively. So, the ratios among the coefficients were equally accurate in the two cases, before and after feedback.

Table 1
Event-Specific Scaling Constants

Event	Physical	From Judgments
FF	76.93	60.54
StS	28.64	23.77
MS	25.21	20.31
ShS	18.05	17.40
SP	16.01	12.18
MP	11.45	9.87
LP	9.24	8.28

Note—FF, free fall; StS, steep slope; MS, medium slope; ShS, shallow slope; SP, short pendulum; MP, medium pendulum; LP, long pendulum.

Table 2
Ratios Between Events for the Scaling Constants
Shown in Table 1

Event	Physical	From Judgments
StS-MS	1.14	1.17
StS-ShS	1.59	1.37
StS-FF	0.37	0.39
StS-MP	2.50	2.41
MS-ShS	1.40	1.17
MS-FF	0.33	0.34
MS-MP	2.20	2.06
ShS-FF	0.24	0.29
ShS-MP	1.58	1.76
FF-MP	6.72	6.13

Note—FF, free fall; StS, steep slope; MS, medium slope; ShS, shallow slope; MP, medium pendulum.

Jacobs and Michaels (2001) argued that the improvement in precision that occurs when observers are given feedback cannot be attributed only to the calibration of scaling coefficients. Overall, precision is a function of both between-observers variations and within-observers variations in judgments. We attributed the reduction in between-observers variation to calibration of scaling coefficients. Jacobs and Michaels showed that reduction in within-observers variation (and presumably, also some reduction in between-observers variability) must be attributed to perceptual learning or attunement. Observers were shown to use feedback to discover which optical variables yield specification of the event properties of interest. A difference in the two cases (i.e., calibration vs. perceptual attunement) is that calibration appears in our study to have occurred very rapidly (within a single trial), whereas perceptual attunement was more gradual, occurring over a significant number of trials, in Jacobs and Michaels’s study. Although a significant proportion of the improvement in performance in our study should accordingly be attributed to calibration (because it occurred within a single trial), some should probably be attributed to perceptual attunement as well.² In fact, we suggest that perceptual attunement should be expected to occur in most perceptual studies. The reason is as follows.

The standard paradigm in perceptual studies is to isolate in displays the information variables to be tested. Isolating hypothesized information variables entails the removal of the structure that the perceptual system may normally rely on. Investigators have effectively assumed that the information variables under study are detected and used independently in normal environments, but this is unlikely to be the case. When the observer is confronted with isolated variables, it is likely that he or she has to learn to use the information in its isolated state. Either observers have to learn to use it for the first time (because they do not normally use it), or they have to learn to use it out of context. A third possibility combines these first two. The structure in context is actually a different (but closely related) variable, and so when taken out of context, the variable is effectively new and must be learned as such. Think of techniques for isolating optic flow from texture gradients. This must be perturbing (and ontolog-

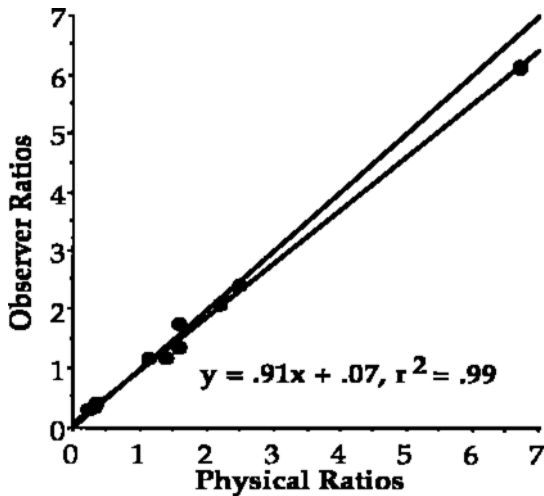


Figure 1. Ratios of mean observer constants regressed against the ratios of physical constants.

ically problematic) because, by definition, optic flow carries optical texture into texture gradients. The two, flow and gradients, are not separate. Similarly, looking at a distant object necessarily entails a number of relational states (accommodation, convergence, parallax, perspective on a ground surface) that did not obtain in our experiments or in Jacobs and Michaels's (2001). Bingham and Shull (1999) found that the presence of a visible ground plane was required to allow calibration of reaches from haptic feedback to generalize to all locations across reach space. So, the isolation of information variables (e.g., removing ground texture to isolate parallax, or event-specific variables) is likely to be highly perturbing and nonrepresentative. The presence of gradual perceptual attunement would provide a good indication that this is so.

REFERENCES

- BINGHAM, G. P., & SHULL, J. A. (1999). Visible support surface allows haptic feedback to generalize across reach space. *Abstracts of the Psychonomic Society*, *4*, 59.
- HECHT, H., KAISER, M. K., & BANKS, M. S. (1996). Gravitational acceleration as a cue for absolute size and distance? *Perception & Psychophysics*, *58*, 1066-1075.
- JACOBS, D. M., & MICHAELS, C. F. (2001). Individual differences and the use of nonspecifying variables in learning to perceive distance and size: Comments on McConnell, Muchisky, and Bingham (1998). *Perception & Psychophysics*, *63*, 563-571.
- McCONNELL, D. S., MUCHISKY, M. M., & BINGHAM, G. P. (1998). The use of time and trajectory forms as visual information about spatial scale in events. *Perception & Psychophysics*, *60*, 1175-1187.
- PIZLO, Z., & ROSENFELD, A. (1992). Recognition of planar shapes from perspective images using contour-based invariants. *Computer Vision, Graphics, & Image Processing: Image Understanding*, *56*, 330-350.
- PIZLO, Z., ROSENFELD, A., & WEISS, I. (1997). The geometry of visual space: About the incompatibility between science and mathematics. *Computer Vision & Image Understanding*, *65*, 425-433.
- RAIBERT, M. H. (1986). *Legged robots that balance*. Cambridge, MA: MIT Press.

NOTES

1. We assumed that observers were not using acceleration. As was discussed by Jacobs and Michaels (2001), Hecht, Kaiser, and Banks (1996) found evidence leading them to suggest that observers use mean velocity, rather than acceleration, to judge spatial scale. This was consistent with other evidence showing that human observers are not very sensitive to accelerations at low to moderate velocities.

2. Some subjects in the experiments of McConnell et al. (1988) were found to be using image size to make their judgments. The displays were designed explicitly to control image size by holding it constant. Accordingly, those subjects were treated as having misunderstood the task and were removed from subsequent analysis. (Their judgments never varied.) Of course, it is precisely this type of performance and subject that Jacobs and Michaels (2001) featured in their study. Because McConnell et al. removed them, the results were bound to be different.

(Manuscript received July 17, 2000;
revision accepted for publication February 14, 2001.)