

Notes and Comment

Support for the cue-heuristic model is based on suboptimal observer performance: Response to Gilden and Proffitt (1994)

SVERKER RUNESON
Uppsala University, Uppsala, Sweden

Gilden and Proffitt (1994) have derived support for their updated cue-heuristic model of mass discrimination in observed collisions from two experiments reported by Runeson and Vedeler (1993). However, these experiments are inadequate for theory testing because of confounding and lack of representative variation in parameters. Instead, extended analyses of the third of Runeson and Vedeler's experiments are presented. It is shown that observers exclusively use collision-axis velocity components in conditions where the heuristic model asserts that only trajectory speeds are available. Moreover, the patterns of error predicted by the heuristic model do not occur. Overall, performance is closer to the ideal than to the predictions from any conceivable model based on elementaristic cues; hence an explanation in terms of advanced informative invariants is the more viable approach. The final discussion concerns the distortion of theory development that might be engendered by empirical data that represent suboptimal observer performance due to unfavorable testing conditions.

In a way potentially relevant for perception, the kinematics of colliding objects contain information that specifies the ratio of the objects' masses and about their efficient elasticity (Runeson, 1977/1983). This finding has been generalized and applied to both animate and inanimate events. Thus the expression *kinematic specification of dynamics*, or the *KSD-principle* for short, is used to identify a category of informative relations, granted by basic natural laws and ecological constraints (Runeson & Frykholm, 1983; see also Runeson, 1994a).

When a certain type of information is shown to exist, it raises the possibility that it might be exploited by per-

ceivers. The prospects are strengthened when, as in the case of KSD, the type of environmental properties specified appears to be of relevance for organisms in planning and performing their actions. Thus it was suggested that, although mass and elasticity are "hidden" dynamic properties of objects, they might nevertheless be perceivable by vision (Runeson, 1977/1983).

In collision events, the information for mass ratio resides in higher order properties (invariants) of the pattern of motions. The informative relation can be described in several ways. For collisions in one dimension,

$$m_B/m_A = |v_A - u_A|/|v_B - u_B|, \quad (1)$$

where m_A and m_B are the masses of the two objects, and u and v designate the velocities before and after impact (Figure 1). For collisions in two and three dimensions, the same equation can be used if the motions are projected on the collision axis. However, one can also use a generalized equation in which motion vectors replace the velocities:

$$m_B/m_A = |v_A - u_A|/|v_B - u_B|. \quad (2)$$

Furthermore, observing that each $(v - u)$ term stands for change in motion in an object, we can set $w = v - u$ and get

$$m_B/m_A = |w_A|/|w_B|, \quad (3)$$

where w_A and w_B are collinear vectors that describe the changes in motion incurred by the objects in the collision. Thus, in the most compact form, the ratio of mass is inversely proportional to, and hence uniquely specified by, *the relative amount of motion change* incurred by the two objects in the collision.

In the following discussion, it is important to consider the ontological implications of this set of alternative equations. Their variety shows that there is no necessity to register vector components of the motions, or to derive them as projections on any explicit or implicit axes. Nor is it necessary to register trajectory parameters (e.g., speeds and directions) or individual motion vectors, inasmuch as a change in motion is equally well a kinematic property. Hypothetically, it is therefore possible that motion change could be picked up directly by a suitably evolved and attuned perceptual system (i.e., a smart perceptual mechanism; Runeson, 1977, 1994b). Indeed, the same possibility exists for the relative amount of change in two moving objects, hence the mass ratio might be directly perceivable as a single property. The plausibility of such a function is augmented by the fact that in collisions the two changes are simultaneous, collinear, and abutted end-to-end.

I am indebted to Thomas Harrington for cooperative efforts on the informational analysis of collision mechanics and to Emily Holmes for help with English. I also thank Geoffrey Bingham, Steven Flynn, Harold Sedgwick, and the Editor for important comments and suggestions. The current research was supported by The Swedish Council for Research in the Humanities and Social Sciences (HSFR), The Bank of Sweden Tercentenary Foundation (RJ), and The Swedish Collegium for Advanced Study in the Social Sciences (SCASSS). Correspondence should be addressed to S. Runeson, Department of Psychology, Box 1854, S-75148 Uppsala, Sweden (e-mail: sverker.runeson@psyk.uu.se).

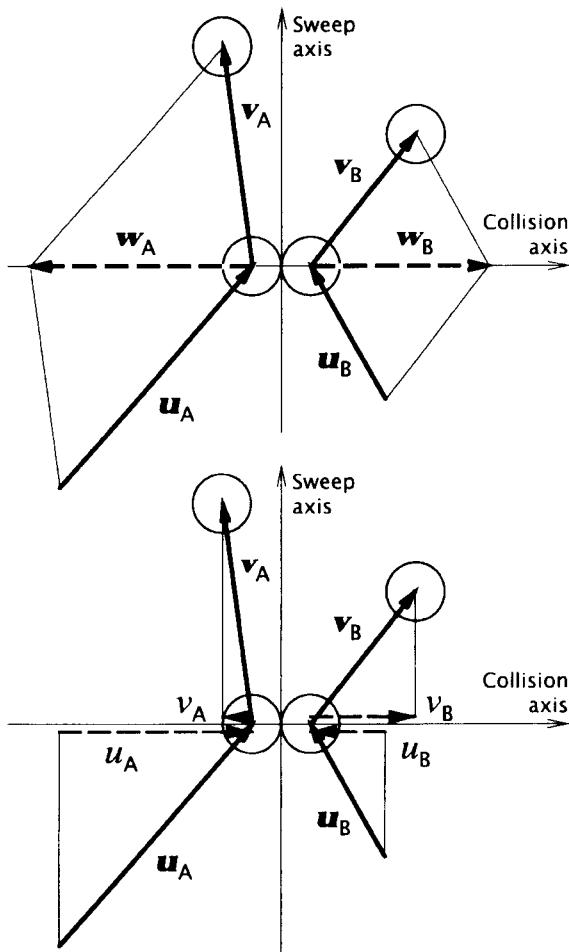


Figure 1. Collision in two dimensions described in terms of velocity vectors u and v for the pre- and postcollision motions. The upper diagram includes the vectors w that describe the changes in motion incurred in the collision. The lower diagram shows the same motion vectors together with their collision-axis components, u and v , which are used to define several key terms. The left-hand object is the *impinging object*, labeled A, because $u_A > -u_B$. This asymmetry in initial motion is called V_{comm} (common-mode velocity). In the experiments discussed, V_{comm} is “small” when $u_A / -u_B = 3$ and “large” when $u_A / u_B = 5$. $V_{\text{diff}} = u_A - u_B$ is the velocity with which the objects strike each other. *Win Velocity* for A is v_A and $-v_B$ for B. *Sweep components* are projections of the motion vectors on the sweep axis. *Exit speeds* are the magnitudes of the v vectors. An object’s *scatter angle* is its change in direction due to the collision. For more details, see Runeson and Vedeler (1993, Figures 2, 5, and 7) and G&P (1994, Figure 4).

Performance Assessment

In a number of studies, researchers have sought to determine empirically whether human observers possess a perceptual skill for mass discrimination in collisions and, if so, to determine its performance characteristics as well as the nature of the underlying psychological process or function. The main papers are Todd and Warren (1982), Gilden and Proffitt (1989), and Runeson and Vedeler (1993). (Henceforth, Gilden and Proffitt will be abbreviated G&P; Runeson and Vedeler, as R&V.) Over

the course of this series of studies, simulation and display techniques have developed, and a more representative sampling of collision parameters has been incorporated in the experimental designs. As it appears, these improvements have allowed progressively more observer skill to become evident in the data.

All studies have employed a binary forced choice task on the quantitative mass ratio variable (“which object is heavier”).¹ Hence the experiments take the form of difference-threshold measurements, and the results can be appropriately described and analyzed by means of the concepts and methods developed in classical psychophysics (Gescheider, 1985; Guilford, 1954). Progressively better results have been reported on both point of subjective equality (PSE) offset (the *constant error*) and consistency (the *variable error*). The most notable departure from ideal performance is a consistent bias: The mass of the impinging object is always overrated to some extent. PSE offsets varying from 54% (Todd & Warren, 1982) and 46% (G&P, 1989) down to 21% (R&V, 1993) have been reported.² This development shows that, under adequate test conditions, observers of collisions can be quite adept at discriminating relative mass. Moreover, a recent study by Flynn (1994) suggests that the PSE offset is even smaller when one is viewing natural rolling collisions, as opposed to simulated sliding collisions. Hence continued research will be required before human performance on this type of task can be fully assessed.

Cues Versus Invariants

Beyond performance characteristics, there are the problems of explanation: How do observers perform the task of discriminating the relative mass of two colliding objects? Over this question a controversy has arisen through G&P’s (1989, 1994; Gilden, 1991; see also Todd & Warren, 1982) taking issue with the approach suggested by Runeson (1977/1983; R&V 1993).

It is characteristic of the mass-discrimination task that there are no *simple cues* that have serviceable correlations with the actual mass ratio. Impotency of simple cues and concomitant availability of specificational invariants is typical of natural instances of perceiving. In any such case, there are two major ways to approach an explanation. The received view entails that perception must make do with simple cues, despite their informational weakness, because they are presumed to be the only kind of properties that are accessible to the visual system—an implicit commitment to elementaristic ontology. Judgments of the target property are therefore held to be achievable only through adjuvant secondary processes: inferencing, integration, heuristics, dominance metrics, and so forth.

In line with this theoretical outlook, G&P (1989) proposed a cue-heuristic model for mass discrimination in collisions. Velocity vector components were recognized as highly informative, yet they were held to be perceptually inaccessible, owing to their status as “multidimensional quantities” (G&P, 1989, p. 373), which would have

to be constructed or derived from speed and direction cues (G&P, 1989, p. 382; 1994, p. 714).³ Only simple trajectory properties were therefore posited as inputs to the judgment process. Supportive evidence was presented. In its current form (G&P, 1994), the cue-heuristic model states that mass-ratio judgments are formed on the basis of three separately perceivable cues: ratio of *exit speeds* and occurrence of *ricochet* (defined as scatter angle $>90^\circ$) in either object.⁴ Moreover, observers do not form their judgments through addition or any other simple operation on the cues. Instead, it is asserted that judgments in each case are based on only one of the cues, the choice of which may depend on salience criteria pertaining to the cues, and on individual predilections among observers.

In this view, "the focus is on the pattern of error." Performance is expected to be fraught with "systematic and egregious errors" (G&P, 1994, p. 708), because the secondary processes are not sufficient to make up for the weaknesses of the inputs. Although observers can make judgments, it is generally denied that dynamic properties are perceivable as such (Gilden, 1991).

The alternative approach maintains that the elementaristic restraint on theorizing is scientifically untenable and counterproductive (Runeson, 1977, 1994b; see also Bunge, 1977, pp. 58, 99; Rosen, 1978; Turvey, 1992). Liberation from elementaristic ontology is characteristic of the Gibsonian ecological approach (e.g., Gibson, 1979). It accepts as a scientific challenge that perceptual systems may function directly on the basis of more advanced properties in proximal patterns. Decent levels of performance may thus be attainable without a fundamental reliance on secondary processing.

Essentially, the question is: Can perception directly take advantage of advanced properties, or is it a priori limited to the elemental properties posited by some variety of physics theory? Research on visual mass discrimination in collisions offers opportunities for empirical illumination of this fundamental issue. Phenomenally, at least in our lab setup, it is a straightforward single-property perceptual task. In most cases there is an immediate and suggestive impression of which object is heavier, while kinematic details such as speeds or directions are noticed only if specifically attended to.

As a first installment in a planned suit of empirical reports, R&V (1993) tested one of G&P's (1989) claims: that observers neither need to use, nor do use, any of the information in the precollision epoch. We found that if a more representative variety of collisions was used, the precollision epoch was indispensable: performance deteriorates drastically if it is not shown. Furthermore, we found evidence that observers do in fact use a kinematic property that is analytically definable in terms of velocity vector components, rather than the less informative exit trajectory speeds.

Application of the Cue-Heuristic Model to Runeson and Vedeler's (1993) Results

In R&V (1993), we discussed the possibility of testing G&P's (1989) heuristic model on our normal-view data.

We found, however, that the model was not specified in a way that allowed its application to the more representative samples of collisions employed in our study. It is therefore helpful that the current version of the model (G&P, 1994) has been improved in this respect. To make the model relevant for collisions in which both objects are in motion before impact, the ricochet and scatter angle of object B are no longer measured relative to the incoming direction of the other object. Instead, they are defined the same way as for object A—that is, with respect to the object's own incoming direction.

Unfortunately, this revision is not sufficient to make the heuristic model fit for detailed predictions. As G&P (1994) underscore (pp. 717, 719), the model will require as yet unavailable knowledge about speed and direction perception, about the salience of perceived speed and angle differences, and about the variety of ways in which observers handle cue conflicts in collision events.

Despite these limitations, G&P (1994) have proceeded to confront their model with the results of our first two experiments. They concluded that a good fit to the data is possible and that the adequacy of the model was verified. To this end they have, first, for each of the main experimental conditions, divided the range of mass ratios into predictive zones delimited by the occurrence of ricochet in either object and by the direction of the exit-speed difference (see their Figures 3 and 8). Then, while considering one or a few data points at a time, they have discussed various assumptions about salience criteria and tradeoffs between cues—for instance, to what extent observers may have been "impressed" by one or the other cue when forming their judgments. As I understand it, G&P's point is to show that *if* the salience limits and the relative impressiveness of the respective cues were adjusted, and *if* observers in some cases were allowed to ignore ricochets and resort to quantitative comparisons of scatter angles, *then* the data points considered would be accounted for by the model.

An ad hoc fit to data that is achieved with several cues, and with salience and tradeoff parameters left open, is usually of limited value. Since parameters cannot be derived from theory, it is therefore unfortunate that G&P have not taken the opportunity to move toward an empirical specification of their model by assigning tentative values to salience limits and tradeoff functions. By doing this, they could have shown that at least the present data sets were predictable and the model would have become potentially useful beyond pointwise explaining. Thus one is left with the question of whether G&P's deliberations over the data at hand could converge to rules that would apply successfully across the range of studied collision parameters.

A closer examination yields negative expectations in this regard. For example, a 50% exit-speed advantage in an object is often designated "large" and "decisive," and 35% leads to "unanimous agreement" (G&P, 1994, pp. 711, 715). Yet, in one case (Experiment 2, $m_B/m_A = 1.81$, V_{comm} large, normal view) G&P face the fact that a 127% speed advantage in object B is not sufficient to

nullify a scatter angle in object A of only 67° , which is 23° below the usual salience limit (p. 717; Figure 7). Unanimous “A” judgments would be strongly expected, yet in a significant 19% of these trials, B was chosen as the heavier. Without further development, G&P comment that “We have seen . . . that the scattering angle may be impressive to observers at this magnitude” (p. 717). Thus, to accommodate this datum, the salience limit for scatter angle needs to be set lower than 90° . It would also be necessary to depreciate the relative impressiveness of the speed cue, even when speed differences are large.

If the model were adjusted in this way, problems would arise at the next larger mass ratios in the same condition, however. In the first of these collisions ($m_B/m_A = 2.69$; see their Figure 7), the scatter angle of A (87°) is substantially above the adjusted salience limit, whereas B’s speed advantage has dropped to an unimpressive 87%. Without the adjustments, this would have been a simple case of cue conflict with a small bias toward “A heavier,” which is reasonably in line with the intermediate response frequency obtained (48%). However, the adjusted model would predict a substantial dominance of “B” responses and be unable to accommodate the empirical outcome.

Finally, at $m_B/m_A = 4$, the scatter angle of A has increased to 101° , and B’s speed advantage is down to 46%, just below the basic 50% salience limit. The unadjusted model would predict a clear dominance of “B” judgments, and the obtained figure, 67%, does not quite bear that out. This limited misfit would get much worse with the adjusted model, however. The scatter cue would dominate completely and unanimous “B heavier” judgments would be predicted.

Thus G&P’s indicated solution to the $m_B/m_A = 1.81$ case, entailing a substantial appreciation of the salience of the scatter cues relative to that of the exit-speed cue, appears to run counter to what is needed at adjacent data points in the same experimental condition. Moreover, it would conflict with the cases in which speed differences in the 35% to 50% range are deemed decisive. And there is no obvious basis in the cue values for why the trade-off function should be different from one data point to the other.

Admittedly, G&P’s model appears to be consistent with the direction and overall magnitudes of the PSE offsets observed in the experiments published up to and including the first two experiments in R&V (1993). This is to assume, however, that only gross matches between theory and data are required—for instance, if threshold crossovers curves are allowed to have any shape and location within the flexible mass ratio ranges where the heuristics are not clearly unanimous. Strictly speaking, this is as much as G&P claim their model can currently do, although their discussion reveals an interest in finer aspects of the experimental results. In all likelihood, the distinguishing of cue-based inferencing from invariant-based perceiving will require closer attention to predictions and data.

Although this state of affairs effectively precludes confirmation of G&P’s model, it is nevertheless possible that it could be empirically defeated, as it were, in its infancy. That is to say, the model may turn out to be so much at odds with empirical data that no future adjustments or specifications could bring it into correspondence. This possibility will be examined in the following.

Most of the amendments to basic rules that G&P discuss concern collisions for which the speed and angle heuristics either conflict or provide no recommendations at all. Hence, as G&P (1994, p. 719) indicate, it should be possible to use the basic rules alone to make global predictions: The mass ratio zones in which response proportions should hover near the 0% or 100% levels could be identified. In between those zones, predictions could be set to 50/50, as G&P do at various points (e.g., G&P, 1994, Figure 9). Alternatively, the intermediate zones can be considered as predictions of the locations and widths of threshold zones, in the sense of classical psychophysics.

G&P (1994) did not apply their model to the results of our third experiment, because the data were not graphed in a way suitable for their kind of interpretation.⁵ Although it would have been preferable had G&P themselves made the predictions (all the necessary information was given in R&V, 1993), the current situation offers an opportunity for confronting their model with relevant data, the details of which were unavailable to G&P when the model underwent revisions. Thus, much of the ad hoc character of the test can be eliminated.

Occluded Precollision: The Speed Heuristic On Its Own

R&V (1993) compared performance in full-view collisions (the *normal* condition) with collisions in which only the postcollision motions were visible (the *occluded* condition). As affirmed by G&P (1994; Gilden, 1991), scatter and ricochet are properties that relate pre- and postcollision motions and are therefore available only when the whole event is visible. Thus they suggest that exit-speed ratio is the only cue on which observers could base their judgments in the occluded condition. Bizarre as these events may appear to observers, the occluded condition therefore offers an opportunity to test the speed heuristic on its own. Hence G&P (1994) made predictions for our Experiment 2, found a close fit with the empirical results (see their Figure 9), and concluded that the exit-speed heuristic was the basis for the judgments made by the observers.⁶

Confounding and its resolution. However, for relevance in the present discussion, G&P’s exit-speed cue must be tested along with the alternative suggested by R&V (1993): *win-velocity difference*. Win velocity was defined as the collision-axis component of an object’s postcollision motion, counted as positive if it “wins” the encounter by proceeding onto the other object’s half of the field (Figure 1). A comparison based on this kinematic property provides the best (i.e., least bad) infor-

mation about mass ratio and is thus the most KSD-like alternative in the aberrant condition of occluded pre-collision (see Figures 7 and 9 in R&V, 1993). However, it turns out that win-velocity difference and exit-speed ratio yield identical predictions for our Experiment 2, so that neither theory can gain support from the results of this experiment.

The reasons for this, discussed in R&V (1993, p. 624 ff.), formed the rationale for our Experiment 3. Exit speed is the magnitude of the vector obtained by adding the orthogonal win-velocity and sweep components. In other words, win velocity and sweep velocity form a right triangle, with exit speed as the hypotenuse (Figure 1). Throughout Experiment 2, both objects had identical sweep components, so there was a unique functional relation between log exit-speed ratio and win-velocity difference. As can be seen in Figure 2, both functions are fairly straight as they pass diagonally through zero at the same mass-ratio values; hence they predict identical PSEs and similar threshold regions. For experimental purposes, they are thus completely confounded.

We therefore designed Experiment 3 in order to unconfound the competing properties. Although collision mechanics and the call for representativity prohibit their independent variation (see Warren & Owen, 1982, for a discussion of the difficulties that this engenders for experimentation and analysis), the functional relation between them was broken up into a dual mapping. The two objects were assigned different sweep velocities.⁷ Hence there were two sweep conditions: either a large sweep component was assigned to object A while object B had a small sweep, or vice versa. These conditions were combined in factorial fashion with the two values on the common-mode velocity variable, forming four experimental conditions. Each contained a run-through of the series of mass ratios tested.⁸ The kinematics of the resulting 32 collisions were illustrated in R&V (1993, Figure 5; see the concluding discussion below for more details on this experiment).

Figure 3 contains separate graphs for each of the main experimental conditions and shows the progression of exit-speed ratio and win-velocity difference as functions of mass ratio. Also shown are the PSEs and widths of the threshold zones predictable from the two candidate properties. With this design, the proposed hypotheses predict different patterns of judgments. Whereas both predict a large effect of the V_{comm} factor, only the speed heuristic predicts an effect of the sweep variable.

Results. Figure 3 also contains the empirical results as actual response proportions and as PROBIT-fitted threshold ogives. By integrating the information in all eight data points, the fitted ogives are superior in describing the empirical threshold characteristics (Guilford, 1954, chap. 6; R&V, 1993, p. 622). For comparison with the predictions, the empirical PSEs are shown with surrounding threshold zones as shaded fields.

The empirical results conform well with the win-velocity predictions and disagree with the predictions from exit speed. The possible exception is the lower right condition, in which the empirical PSE falls midway between the two predictions. However, in this condition as well as the one above, the threshold analysis builds on extrapolation beyond the studied range of mass ratios. Hence it is more sensitive to fortuitous variations in the rightmost data point.

A close examination of the relations between the candidate properties and the data within the studied range of mass ratios reveals that the discrepancies for exit-speed ratio are large. Further, the differences in fit are not just quantitative. Data exhibit a qualitative pattern predicted only by the win-velocity property: invariance over the sweep conditions.

In conformity with G&P's point-by-point approach to theory testing, the analysis entailed in Figure 3 was done in four parts, one for each of the main experimental conditions. Although the failure of the exit-speed heuristic is shown beyond doubt, it is difficult to gauge the overall difference in explanatory value between the two can-

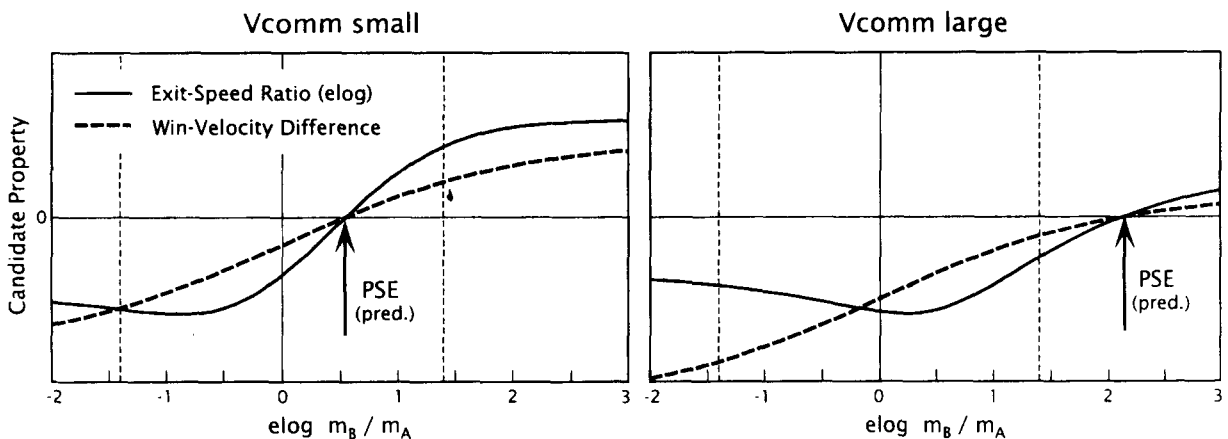


Figure 2. Progression of exit-speed ratio and win-velocity difference under the conditions used in Runeson and Vedeler's (1993) Experiment 2 (vertical scales are arbitrary; cf. Figure 3). For clarity, the functions are shown over an extended range of mass ratios; dashed vertical lines delimit the studied range.

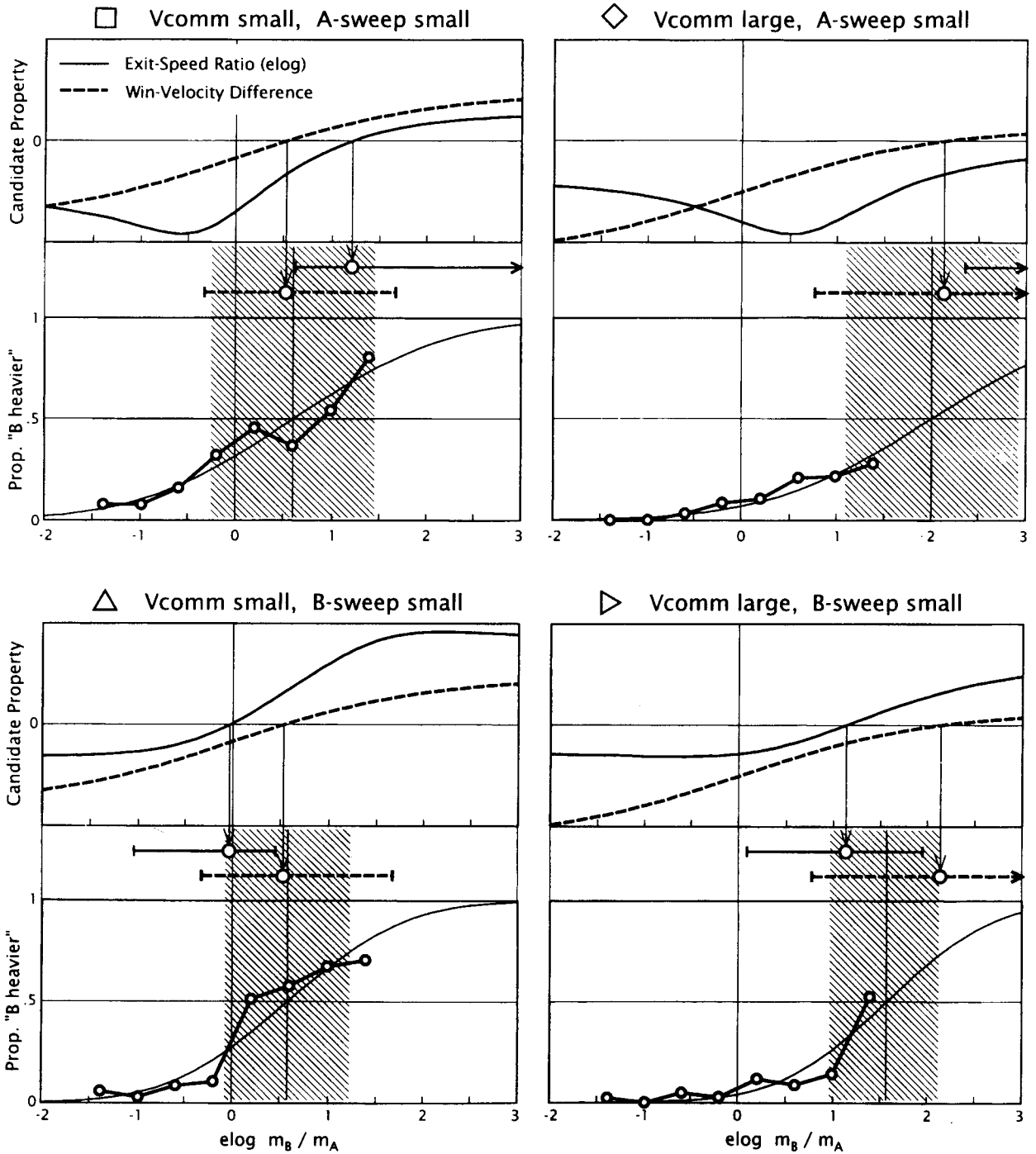


Figure 3. Predictions and results for the *occluded* precollision condition in Runeson and Vedeler's (1993) Experiment 3. The four experimental (initial-motion) conditions are treated in separate panels. Panel headings include the symbols used to identify the data points in Figure 4. Upper subpanels show the two candidate kinematic properties, exit-speed ratio (solid lines; $e \log$ range = ± 1.5) and win-velocity difference (broken lines; range = $\pm 3 \cdot V/diff$), as functions of mass ratio. Threshold predictions are shown as horizontal bars in corresponding line styles. Circles mark predicted PSE points; arrowhead ends indicate predicted threshold zones extending outside the graph range. PSE points are predicted where the respective candidate property crosses zero. Uncertainty zone limits are set at $2/3$ and $3/2$ for exit-speed ratio ($\pm 50\%$ speed difference), and at $\pm 8 \cdot V/diff$ for win-velocity difference. Lower subpanels present the empirical results as response proportions and as fitted threshold ogives, partly extrapolated. The shaded rectangles describe the width (proportion = $.5 \pm .25$) of the empirical threshold zones, each divided by a vertical line at the estimated PSE.

didate properties. The hallmark of a perceptually effective property is *invariance of its fit to data over all experimental conditions*. Hence, the numeric value by which it describes each collision should rank all of the 32 different collisions used in the experiment in the same way that the observers do. When a two-alternative discrimination task is used, the functional relation between the effective property and the response proportions is ideally a step function. However, because of inescapable noise, or momentary fluctuations in the threshold, the outcome should conform to a cumulative normal ogive.⁹ If two collisions have the same value on the effective property, they should yield the same proportion of "B heavier" judgments even if they differ in mass ratio or belong to different experimental conditions.

The degree of data orderliness brought forth by candidate properties can be described in a very succinct way, as we reported previously (R&V, 1993, Table 4 and Figure 8). By means of modified PROBIT analyses, the fit to data of four candidates, including exit-speed ratio and win-velocity difference, was assessed. A goodness-of-fit measure (McFadden's R^2) was obtained for each candidate. The outcome was very clear: While win velocity scored .42, exit-speed ratio scored only .25.

The attendant graphs are especially telling (R&V, 1993, Figure 8). For further clarity, two of them are repeated here in augmented form, and a comparison with mass ratio is also included (Figure 4).¹⁰ The four experimental conditions are distinguished by different symbols, so that one can trace how the data points become organized differently along the horizontal axis. Plotted against mass ratio, the judgments are highly disorganized—which is not surprising, since the information in the precollision epoch was not available. For win velocity alone, the horizontal shifting makes the data points rally to form a steep normal ogive that identifies the perceptually effective property. Exit speed exhibits an intermediate fit, not because observers have used it, but as an artifact of the moderate correlation that it has with win velocity.

In conclusion, the only available empirical material that is relevant to the exit-speed versus win-velocity contrast is the occluded-precollision part of our Experiment 3. Clearly, it does not support "the reality of the speed heuristic" (G&P, 1994, p. 719). On the contrary, it demonstrates the perceptual efficiency of the more informative win-velocity property. It follows that the fit of exit speed to our Experiment 2 data, presented by G&P (1994) in their Figure 9, is incidental and due to the confounding inherent in the design of that experiment.

Normal-View Collisions

Although tested only with occluded precollision data, the nonappearance of the exit-speed heuristic and the concomitant demonstration of observers' ability to use a velocity component property make it difficult to see how the cue-heuristic model could be retained. Possibly, one could argue that the occluded condition is exceptional in this respect. With normally viewed collisions, the exit-speed cue could—not by any known principle, but as an empirical fact—turn out to be effective, together with ricochets or scatter angles. For this reason, G&P's model should also be confronted with normal-view data.

The confounding between exit speeds and collision-axis velocity components that occurs in our Experiment 2 also curtails its validity for the present purpose. It therefore remains to test the predictions possible with the current version of G&P's model against the results of our Experiment 3.

Results. Figure 5 contains an array of graphs that describe the progression of the exit-speed ratio and scatter-angle cues over the range of mass ratios, together with predictions and the empirical data. The cue-based predictions were obtained by applying basic salience criteria and combination rules as they appear in G&P (1994). Details are described in the figure legend.

The graphs illustrate the informational weakness of the cues, owing to their variform relations to mass ratio

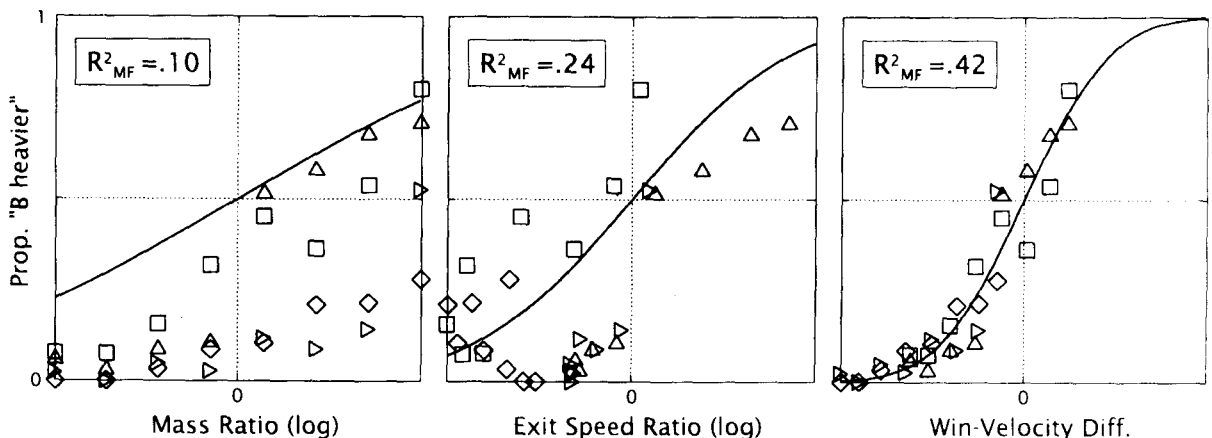


Figure 4. PROBIT analyses of the *occluded-precollision* results of Experiment 3 in Runeson and Védeler (1993). The symbols distinguish the series of data points that belong to the four main experimental conditions, as identified in Figure 3.

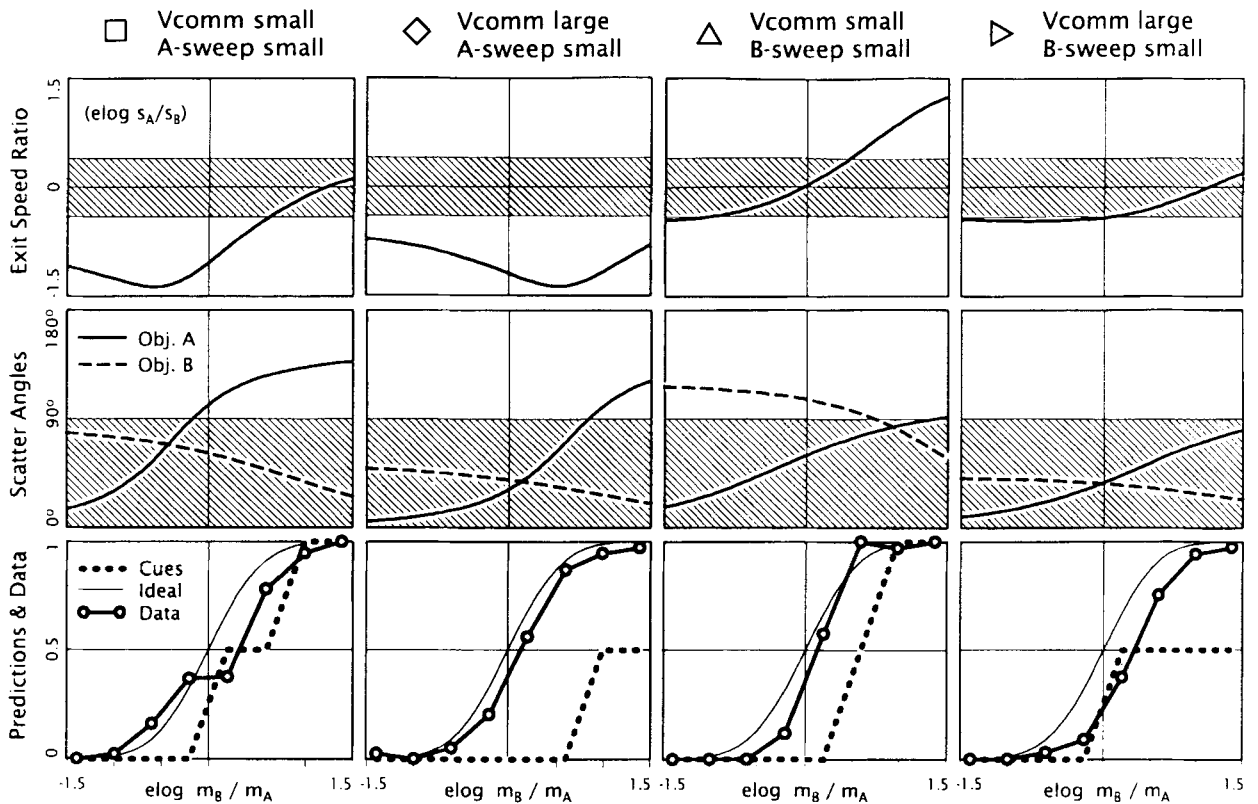


Figure 5. Elemental cues, predictions, and empirical results from Experiment 3 in Runeson and Vedeler (1993), *normal-view* condition. Each column of graphs pertains to one of the experimental conditions. The top two rows depict the exit-speed ratio and scatter-angle cues as functions of simulated mass ratio. The bottom row contains both predictions and the response data (proportion “B heavier”). Shaded areas indicate the values outside of which the cues are salient. The criterion for exit-speed salience is a 50% speed advantage in either object. Scatter angles become salient when “ricochet” begins (i.e., when they exceed 90°). Cue-based predictions are set to .5 when cues conflict or provide no recommendation. The function described by the ideal KSD-invariant is not shown, since it would simply consist of a diagonal line through the center of each graph. The slope of the ideal prediction ogives is not derivable from theory and hence has been set to the average slope of the data curves.

in a more representative sample of collisions. The sensitivity of the cue functions to the sweep conditions should be particularly noted, since this is how unconfounding was achieved. Our Experiment 2 included only two sets of exit-speed and scatter functions, each one intermediate between the current first and third columns and the second and fourth columns, respectively. Thus, far from invariance, both the exit-speed ratio and the scatter angles are strongly interpenetrated by the circumstantial variables: common-mode velocity and the sweep configurations. For example, when an object is given a larger sweep component, its exit speed increases and hence it should look lighter. At the same time its scatter angle gets reduced, which should make it look heavier. Although the effects are in opposite directions, the relations are complicated and they do not cancel out, except for occasional parameter constellations.

This analysis substantiates G&P’s contention that, according to their model, performance on a relative-mass discrimination task should vary greatly between experimental conditions. Even in normal-view collisions, the model predicts large effects of both *V*comm and sweep

variations, whereas the current results show relatively minor effects of these variables.

Generally, the cue-heuristic model predicts bad performance in the form of large PSE offsets to the right in the graphs. In the actual performance, the offsets are much smaller; in fact the empirical data are closer to the ideal than to the cue-based predictions. Moreover, the empirical threshold zones extend quite far to the left. In most of those cases none of the cues could support “B heavier” judgments. Hence the substantial proportions of “B heavier” responses that make up the left-hand sides of the distributions cannot be explained in the way that G&P suggest: They cannot have resulted from individual preferences for either one or the other heuristic among the observers.

The comparisons in Figure 5 were based on a certain way of combining the cues, inspired by, but not strictly derivable from, G&P (1994). Thus one might contemplate other ways of defining or combining the cues that might yield better predictions. For instance, it might be possible to improve predictions in one experimental condition by dropping or lowering the 50% salience limit on

exit-speed ratio. However, even in this case, it appears that it would be difficult to find revisions that would not make predictions worse when applied to the other conditions. Thus no decisive improvement in overall model fit seems to be possible in this fashion.

Since G&P's model does not allow quantitative predictions in the normal-view condition, it is not possible to use the PROBIT method for an overall evaluation of the model as such. Nevertheless, the method can be used to illuminate the explanatory potential of the cues involved. To do so, the two binary scatter cues (A-ricochet and B-ricochet) are treated as a single quantitative cue: *scatter-angle difference*. The results are shown in Figure 6.

The outcome is clear-cut: Exit-speed ratio has no explanatory value, whereas scatter-angle difference fares a good deal better. This is not surprising; scatter angles are defined over both epochs of the event, whereas exit speeds are not. It is inconceivable that the degree of fit exhibited by the scatter-angle difference property alone could be significantly improved by reducing it to ricochets or by conjoining it with exit-speed ratio. Indeed, the analysis suggests that a better model might be achieved by dropping both exit speeds and ricochets entirely and positing scatter-angle difference as the only cue. Although such a revision would require abandonment of the more-than-one-heuristic notion, it would remain true to G&P's fundamental postulate about exit speeds and scatter angles as the only kinematic properties that are perceptually accessible.

Cues aside, the most important feature of the results revealed by Figure 6 is that the best fit to the judgments is achieved by the target variable, mass ratio. This is to say that the correct KSD invariant, expressed in suitable form, is the closest that we can currently get in describing the information used by observers in visual discrimination of the relative mass of colliding objects.

Concluding Discussion

Experiment 3 in R&V (1993) was designed for more limited purposes than a general evaluation of the com-

peting approaches. It therefore has several features (e.g., the inclusion of occluded-precollision blocks, the sweep-glitch manipulations, the omission of practice blocks in the normal-view condition) that may have interfered with performance. For instance, it may have led to less steep threshold ogives.

Despite its limitations, our third experiment provides the current benchmark for human performance on mass discrimination in simulated sliding collisions. Overall, the PSE occurs when B is 21% heavier than A; this is less than half the offset found in other experiments. Furthermore, it includes the most nearly representative sampling of the relevant collision parameters available so far and is thus the only experiment in which trajectory speeds and collision-axis velocity components are usefully unconfounded. For these reasons, it is currently the only published experiment that allows meaningful comparisons of the competing theories. The support that G&P (1989, 1994) have derived from previous experiments for their cue-heuristic model is therefore invalid.

The evidence on perceptual functioning gained from our Experiment 3, amplified by the extended analyses presented here, stands in contrast to the predictions and general expectations derivable from G&P's model. Moreover, the potential of their theoretical approach to the perception of dynamic properties is undermined, because its founding elementaristic assumption ("trajectory properties only") is among the items violated by the empirical results.

Instead, the evidence favors the ecological KSD-based approach with its expectation that mass ratio can be discerned quite well in collisions through the pickup of advanced, informative kinematic properties (Runeson, 1977/1983). It seems fairly clear that observers are able to use information of higher adequacy than the elemental cues suggested by G&P. Since performance exhibits a fair amount of invariance over circumstantial conditions, we are well supported in concluding that relative mass is perceivable by vision. Hence, a clear counterexample is provided to Gildea's (1991) negative assertion concerning the perceptibility of dynamic properties.

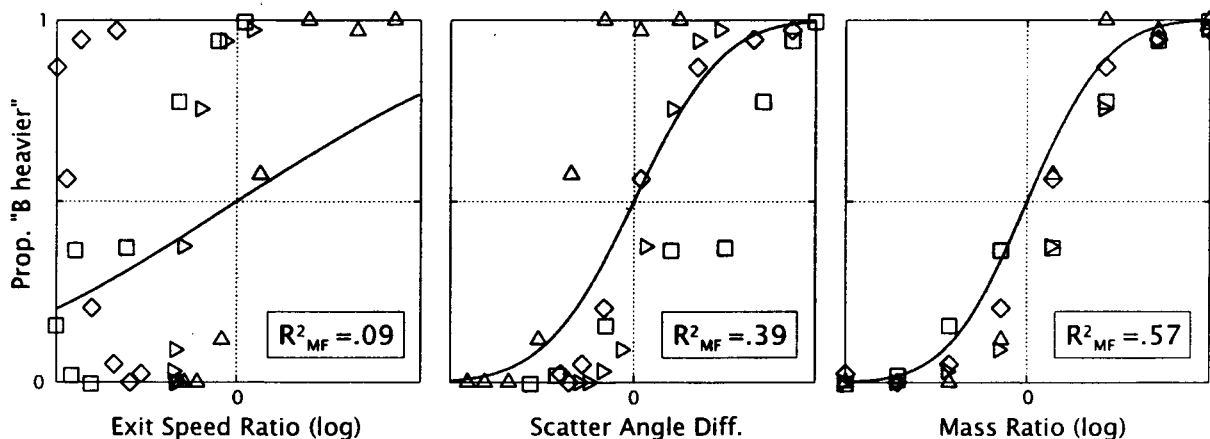


Figure 6. PROBIT analyses of the *normal-view* results of Experiment 3 in Runeson and Vedeler (1993). The symbols distinguish the four main experimental conditions, as identified in Figure 5.

Perceptual imperfections. Although performance surpasses all cue-heuristic expectations, the current results in the normal-view condition still exhibit certain deviations from the ideal, both systematic and variable. Since the overall results are relatively close to the ideal, the deviations are more properly considered as imperfections in the use of a KSD-invariant than as indications of an underlying play of diverse heuristic inferencing, shaping the data set.

Several approaches might be worth pursuing in the search for explanations of the imperfections. First, the PSE offsets may result from *idiosyncrasies* in the operating characteristics of the visual system. For instance, there might be some slippage in the way vector components of motions are picked up, which in turn may occur in the formation of configurations among moving elements, as studied by Johansson (1950/1994). Second, and overlapping with the previous explanation, the kinematic property picked up might be slightly different from an ideal KSD-invariant. It might be what I have called an “incomplete invariant” (Runeson, 1989a, 1989b), or a *nonelementaristic heuristic* in the sense of Braunstein (1976, 1994). Third, the “heaviness” property, as perceptually identified, might not correspond strictly to the property of “mass” as defined in physics. In line with findings concerning the perceptual meaning of “velocity” (Bingham & Runeson, 1983; Runeson, 1974, 1975, 1994b), perceived heaviness might to some extent be infused with the “effort” or “push” exerted in a collision, thus possibly identifying a more action-relevant property. Fourth, observers may be resorting to conscious inferencing in the central zone where mass ratio is too close to unity to be perceptually discerned.

It should be noted that these possible explanations leave room for individual differences between observers and between experimental procedures. Thus they might also come to explain some of the relatively larger constant and variable errors that have appeared in some results, including bimodalities of the sort reported by G&P (1989).

Our Experiment 3 differed from previous research in that somewhat experienced observers participated. Thus, it might be speculated that beginners may function according to the cue-heuristic model and get to use advanced information only as they become experienced in performing the task. However, there are no data that clearly suggest this. Although the early experiments, including our Experiment 2, resulted in large V_{comm} effects, it is not clear how much of the poor performance was attributable to the observers’ lack of experience and how much was due to unfavorable testing conditions. Importantly, no data are available on novice performance under unconfounded conditions. Thus, we do not know whether such data would also exhibit the sweep effects predicted by the heuristic model (see Figure 5).

Generally, the quest for even higher skill levels should continue, and the role and nature of skill acquisition on the task should be studied. More experiments that extend the sampling of the relevant parameter space are also

needed. This will include going beyond the realm of simulated sliding collisions and into more natural conditions, as initiated by Flynn (1994). Thus, the current situation offers a rich variety of topics for further research, many of which could be of relevance for central issues of perception and cognition in general.

Theory development and suboptimal performance data. As G&P agree, mass discrimination in observed collisions has proved to be a useful task for critical evaluation of cue-based versus invariant-based theories of perception. This is due to the higher complexity of this task in comparison with tasks more typically used in perception or information-processing research. The crucial difference, however, is that the informative relations between proximal and distal properties arise through the operation of natural laws and constraints and are not made up by an experimenter’s conjoining of unrelated elements or properties into a synthetic task. With naturally constrained complexity, perceptually relevant higher order information becomes available (Runeson, 1988) and simple cues get insufficient. It follows that a higher task complexity should bring about declining performance if perception is based on cues. Conversely, performance should not be so affected if perception relies on advanced informative properties. Thus, distinct predictions are generated.

Through a long history, cue-based theorizing has become firmly ingrained in perception and cognition research. Hence it is only natural that it gets tried when the existence of a new domain of perceptual achievement is suggested. As G&P have pointed out, the sizable variable-error distributions are especially suggestive of statistical covers under which one might find observers struggling with a variety of heuristic rules.

What is puzzling, however, is that a cue-based model seemed for so long to be empirically vindicated. As we have seen above, one reason was that the first several sets of results, including R&V’s (1993) first experiments, were marked by confounding and were lacking in parametrical representativity. In addition, as we now know, these experiments exhibited levels of performance that were not representative of actual human ability. The large errors exhibited by the results let theorists retain the hope that a cue-based model could be shaped to fit those data. In other words, it appears that perceivers had not yet had their full say on the nature of their perceptual skills. When they were granted better opportunities for that, as in our Experiment 3, their performance exceeded what could be achieved through processing of the cues that constitute the inputs to G&P’s model.

For perception and cognition research in general, it is noteworthy that the cue-based style of theorizing exhibits a certain lack of inherent correctives for ineffectual experiments because of its emphasis on weak and irregular performance. Thus, an experiment that is unsuitable in design or procedure could easily provide supportive-looking data—at least as long as cues, salience limits, and trade-off functions are not specified in detail. This would naturally make it harder to maintain the incentive to look

self-critically for possible reasons why an experiment might have failed to detect a higher level of performance that observers might actually be capable of.

With an invariant-based approach, the incentives are reversed. Thus, the experiments reported in R&V (1993), in addition to testing the role of the precollision epoch, took the form of a search for a methodology that might reveal higher levels of skill, and eventually met with some success in this respect. Flynn's (1994) study is another example that extends the search for true perceptual skill levels to more natural situations.

Given the present examples of how data that represent suboptimal observer performance seem to have provided spurious support for untenable theoretical commitments, it becomes an urgent and delicate concern to find out what the crucial differences between experiments are. As a further example of performance differences, consider the shape of the threshold curves obtained by G&P (1989) and by R&V (1993). In our study, the response proportions for collisions on either side of the crossover region of mass ratios were 0% or 100%, with few exceptions. In contrast, the data curves in G&P's first experiment seem to approach asymptotes at about 6% and 15% (see G&P, 1994, Figure 1). This remarkable difference in data noise levels suggests that experiments can differ extensively in performance-relevant respects and thus give rise to unfair portrayals of observers' perceptual skills.

It follows that methodologies must be critically examined, both in their technical and procedural aspects. Unfortunately, we must be prepared for painful revisions of common practices in experimental work. For the readers' consideration, and to conclude this comment, I will thus offer a list of features of the third experiment in R&V (1993) whereby that experiment differs from some or all of the previous research, features that may have been conducive in precipitating a higher level of performance. In all likelihood, there are even better skills to be discovered if continued attention is given to the methodological aspects.

Simulation technique. The collisions were simulated with an analog computer. The resulting motions were perfectly continuous and thus free of the conspicuous digitization jerkiness that occurs with customary digital simulations. Moreover, analog simulation occurs in real time, so there can be no temporal glitches.

Display. Observers were not positioned to look directly at the oscilloscope. Instead, its image was projected onto a back-projection screen and stray light was added. This in turn was observed through a large collimator lens. In this way, the superficiality, and hence the apparent immateriality, of the displayed objects was reduced, and observers had the impression of looking into a field in which veritable events took place.

Experimental design. In addition to being a richer sample of the collision parameter space, the collisions were displayed in 24 randomly varied orientations. This may have helped observers to bypass simple cues and

urged them to discover invariants. In addition, it also prevented recognition of individual collisions.

Tempo. Care was taken to pace the observers at a comfortable, yet invigorating, rate by adjusting the intervals between the phases of the trials. The time-scale of the simulations—that is, the incoming velocities—was chosen so that the events seemed neither hurried nor sluggish. Pauses were automatically suggested at regular intervals and, in addition, they could be initiated by the observers at any time.

Disturbing factors. It is infeasible to always have both objects start outside the viewing area. An instant onset of motion in a visible object becomes a dynamic event in itself, which, by way of KSD, contradicts the materiality of the objects. Through its abruptness, such a start may also distract the observers' attention from the subsequent collision. Thus the objects were always given a simulated *natural start*, which has been shown to be of minimal conspicuousness (Runeson, 1974, 1975, 1977/1983). Furthermore, the normal and the occluded collisions were placed in separate blocks, following repeated hints that they were rather different types of tasks requiring different modes of attention. Thus a harassment engendered for observers by the random intermixture of occluded and normal-view trials in the first experiments was removed.

Observer experience. All observers, which included the two authors, had experience with normal-view collisions from at least one other experiment in our series and were medium or high scorers.

Motivation. Instructions were given with care. The observers were told that the simulations were physically correct and that one of the objects was in fact always heavier. It was explained that the study concerned their peak discriminative skill, and they were urged to take breaks whenever they felt their attention was slacking. Observers were paid in proportion to their scores.

REFERENCES

- AMEMIYA, T. (1981). Qualitative response models: A survey. *Journal of Economic Literature*, **19**, 1483-1536.
- BINGHAM, G. P., & RUNESON, S. (1983, October). *On describing what is perceived: Seeing "velocity" vs. seeing "push" in moving objects*. Paper presented at the Autumn Meeting of the International Society for Ecological Psychology, Trinity College, Hartford, CT.
- BRAUNSTEIN, M. L. (1976). *Depth perception through motion*. New York: Academic Press.
- BRAUNSTEIN, M. L. (1994). Decoding principles, heuristics and inference in visual perception. In G. Jansson, S. S. Bergström, & W. Epstein (Eds.), *Perceiving events and objects* (pp. 436-446). Hillsdale, NJ: Erlbaum.
- BUNGE, M. (1977). *Treatise on basic philosophy: Ontology I. The furniture of the world*. Dordrecht: Reidel.
- FLYNN, S. B. (1994). The perception of relative mass in physical collisions. *Ecological Psychology*, **6**, 185-204.
- GESCHEIDER, G. A. (1985). *Psychophysics: Method, theory, and application* (2nd ed.). Hillsdale, NJ: Erlbaum.
- GIBSON, J. J. (1979). *The ecological approach to visual perception*. Boston: Houghton Mifflin.
- GILDEN, D. L. (1991). On the origin of dynamical awareness. *Psychological Review*, **98**, 554-568.

- GILDEN, D. L., & PROFFITT, D. R. (1989). Understanding collision dynamics. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 372-383.
- GILDEN, D. L., & PROFFITT, D. R. (1994). Heuristic judgment of mass ratio in two-body collisions. *Perception & Psychophysics*, **56**, 708-720.
- GUILFORD, J. P. (1954). *Psychometric methods*. New York: McGraw-Hill.
- JOHANSSON, G. (1994). Configurations in event perception. In G. Jansson, S. S. Bergström, & W. Epstein (Eds.), *Perceiving events and objects* (pp. 29-122). Hillsdale, NJ: Erlbaum. (Original work published 1950)
- PROFFITT, D. R., & GILDEN, D. L. (1989). Understanding natural dynamics. *Journal of Experimental Psychology: Human Perception & Performance*, **15**, 384-393.
- ROSEN, R. (1978). *Fundamentals of measurement and representation of natural systems*. New York: North-Holland.
- RUNESON, S. (1974). Constant velocity—Not perceived as such. *Psychological Research*, **37**, 3-23.
- RUNESON, S. (1975). Visual prediction of collision with natural and nonnatural motion functions. *Perception & Psychophysics*, **18**, 261-266.
- RUNESON, S. (1977). On the possibility of "smart" perceptual mechanisms. *Scandinavian Journal of Psychology*, **18**, 172-179.
- RUNESON, S. (1983). *On visual perception of dynamic events* (Acta Universitatis Upsaliensis: Studia Psychologica Upsaliensia, Serial No. 9). Stockholm: Almqvist & Wiksell. (Original work published 1977)
- RUNESON, S. (1988). The distorted room illusion, equivalent configurations, and the specificity of static optic arrays. *Journal of Experimental Psychology: Human Perception & Performance*, **14**, 295-304.
- RUNESON, S. (1989a). An invariant-based model for the acquisition of perceptual skills. In *Coupling perception and action: Proceedings of the Fifth International Conference on Event Perception and Action* (p. 143). Oxford, OH: Miami University.
- RUNESON, S. (1989b). A note on the utility of ecologically incomplete invariants. *International Society for Ecological Psychology Newsletter*, **4**(1), 6-9.
- RUNESON, S. (1994a). Perception of biological motion: The KSD-principle and the implications of a distal versus proximal approach. In G. Jansson, S. S. Bergström, & W. Epstein (Eds.), *Perceiving events and objects* (pp. 383-405). Hillsdale, NJ: Erlbaum.
- RUNESON, S. (1994b). Psychophysics: The failure of an elementaristic dream. *Behavioral & Brain Sciences*, **17**, 761-763.
- RUNESON, S., & FRYKHOLM, G. (1983). Kinematic specification of dynamics as an informational basis for person and action perception: Expectation, gender recognition, and deceptive intention. *Journal of Experimental Psychology: General*, **112**, 585-615.
- RUNESON, S., & VEDELER, D. (1993). The indispensability of precollision kinematics in the visual perception of relative mass. *Perception & Psychophysics*, **53**, 617-632.
- TODD, J. T., & WARREN, W. H. (1982). Visual perception of relative mass in dynamic events. *Perception*, **11**, 325-335.
- TURVEY, M. T. (1992). Affordances and prospective control: An outline of the ontology. *Ecological Psychology*, **4**, 173-187.
- WARREN, R., & OWEN, D. H. (1982). Functional optical invariants: A new methodology for aviation research. *Aviation, Space, & Environmental Medicine*, **53**, 977-983.
2. These data are for asymmetric high-elasticity collisions ($e = 1.0$ or .9). PSE values were obtained by interpolation from Todd and Warren's (1982) Table 2, and from G&P's (1994) Figure 1, no-occlusion condition. The third value is a PROBIT-fitted PSE for all the normal-view conditions in R&V's (1993) Experiment 3. It should be noted that the performance level exhibited in the latter experiment has been replicated in other experiments (to be reported).
3. The elementaristic alternative may be seconded by observations of how people reason about physics problems; see G&P (1989), Proffitt & Gildea (1989).
4. If we consider that two colliding objects can ricochet somewhat independently of each other, it is more correct to say that G&P's model includes three cues. Also, it is noteworthy that the model does not include the incoming speeds as cues, following a previous theoretical commitment to simultaneously available quantities (G&P, 1989, p. 374).
5. G&P's problems pertain to our Figure 6 (R&V, 1993), which was intended to show the effects of the sweep and sweep-glitch factors; hence, the data were collapsed over the common-mode velocity variable. Self-evidently, this variable was not collapsed in our analyses (e.g., Table 4; Figure 8); thus, G&P's critical remarks on this are unwarranted.
6. G&P (1994) also discussed our Experiment 1 data (R&V, 1993). That experiment was a partial replication of G&P's (1989) Experiment 1. As we have explained (pp. 618 ff., 623), these results are irrelevant because of lack of variation in crucial parameters.
7. While the A:B sweep velocity components were 1:1 in Experiment 2, the ratios used in Experiment 3 were 2:5 and 5:2 (R&V, 1993, pp. 623, 626). Thus, G&P's description of our unconfounding procedure as "minor adjustments to the sweep components" (G&P, 1994, p. 718) is incorrect.
8. In addition, our Experiment 3 contained a sweep glitch variable intended to upset specification of the orientation of the collision axis. No effect of it was found, however, and possible reasons for this were discussed (R&V, 1993, p. 627). For instance, the glitch manipulation was sometimes detectable as such and the expected effects were small in relation to the large PSE offsets obtained. Since the glitch factor did not appear to be relevant for the present discussion, the data are collapsed over this variable throughout this paper.
9. G&P's (1994, pp. 716 ff.) discussion does not observe these basic notions, which are common to binary-response methodology (e.g., classical psychophysics and PROBIT analysis; Amemiya, 1981; Gescheider, 1985; Guilford, 1954). Thus, G&P take the sometimes steep data curves as evidence for the hypothesized categorical nature of the underlying heuristic process. If observers could perceive mass ratio as a continuously variable property, G&P argue, the data curves would be more continuous across the mass ratio range. It is notable, first, that there is only one curve that has the affinity of a step function (G&P, 1994, Figure 5). Second, their argument appears to undermine the relevance of the graded dissimilarity ratings obtained in one of their own experiments (G&P, 1989, Experiment 2), the results of which are repeatedly cited as major evidence for their model (e.g., Gildea, 1991; G&P, 1994). Apart from that, G&P's argument is invalid because it disregards that steep data curves should be a natural consequence of the task given to the observers—namely, to give dichotomizing judgments on a variable (perceived mass ratio) that may very well be continuous in experience. G&P also overlook that an artifactual discretization of the data plots occurs if an experiment samples the judged variable too sparsely in relation to the observers' discriminative acuity.
10. Since the publication of R&V (1993), an error in our data-handling program has been corrected. As a consequence, the data points shown in the figures of the present paper are not numerically identical with those in R&V (1993). However, the differences are small and unsystematic, and no revisions in conclusions are indicated.

NOTES

1. The exception is G&P's (1989) second experiment, in which observers also gave quantitative ratings of dissimilarity. However, G&P did not report how the ratings were related to the target variable (mass ratio), only how they were related to certain cue constellations. Hence performance cannot be evaluated.

(Manuscript received December 12, 1994;
revision accepted for publication May 10, 1995).