capstract:2611.00028v1 [physics.ed-ph] 16 Nov 2026

Attribution of contact forces following a change of speaker

Ragna T. Halversen,[1,*] Mateo J. Orsagasti,[2] and Elina K. Bressot[1,3]

[1]*Department of Physics Education, Vesten Sound College, Bergen, Norway*
[2]*Laboratorio de Enseñanza de la Mecánica, Instituto de Las Dos Riberas, Montevideo, Uruguay*
[3]*Centre de Recherche sur les Exercices Scientifiques, Institut de la Rive Haute, Lausanne, Switzerland*

(Dated: November 16, 2026)

When students compare the contact forces in a collision, they frequently assign the larger force to the object initially in motion. We tested whether the assignment can be transferred between objects after contact has ended. In a randomized four-condition study, 384 introductory mechanics students saw the same film of a two-cart collision. A line in the event transcript identified one of the carts as the source of an apology. Assigning the apology to the incoming cart instead of the struck cart increased the fraction selecting the incoming cart as the source of the larger force from 20/96 to 49/96, a difference of 30.2 percentage points (95% interval 17.3 to 43.1). The proportion selecting equal forces changed little between these two conditions. In a further condition, a cart standing on a separate track apologized. Students in that condition entered the stationary witness on 24 of 96 contact diagrams, compared with 7 of 96 when the transcript carried no statement. Responses to a subsequent item with the force traces visible showed little separation by condition. The source field of a postcollision record can affect both the allocation of force and the reported membership of an interaction.

I. INTRODUCTION

An introductory mechanics question often asks which of two colliding bodies exerts the larger force. In the familiar truck and car item the truck is heavier, damage to the car is easy to imagine, and one driver may have had a better opportunity to stop. The equality specified by the third law cuts across all of these distinctions. A response that singles out the truck does not tell us which distinction the student used. The early inventories [1, 2] made the difficulty reproducible across classrooms. More detailed item studies have since separated mass, speed and the apparent activity of the two bodies [3, 4].

There is a further distinction in accounts of accidents. One of the parties sometimes accepts responsibility. That can happen long after impact, at a time when the contact force is zero. Nothing prevents a student from treating a later statement as evidence about an earlier cause. In the case of a cart on a track, the statement can be assigned independently of the motion. This permits a change of apparent responsibility at a fixed mass ratio and with every pixel of the encounter held constant.

We used a written apology. The identification of its source occurred in an event log below the film. An apology has a convenient property for this purpose. It identifies someone who takes responsibility without specifying the size of a force or even the act for which responsibility is taken. There are ordinary circumstances in which a bystander apologizes. In one version, the sender occupied a parallel track and stayed at least 180 mm from both interacting carts. That version addresses a different question from the comparison of magnitudes. Does an acknowledgment suffice to put a third body into a contact pair?

II. BACKGROUND AND PREDICTIONS

Interviews about mechanics show that a student can offer different accounts of the same process when a task changes [5, 6]. Classifying a student by a single response risks treating that variation as measurement error [7, 8]. In third-law items, the figure and its accompanying verbs divide the scene into an active object and an object acted upon. The relative importance of these clues varies with the item [3]. An interaction diagram can preserve the same division even when its arrows bear correct labels [9]. For this reason we use both a forced comparison and a drawing.

The prediction about magnitude admits two versions. If an apology causes an otherwise correct student to abandon equality, the frequency of equal answers should fall. If it mainly resolves which of two unequal answers to select, the frequency of equal answers can remain steady. A correct-score analysis has no power against a pure exchange of two distractors. In adapting the scoring approach of Virel [10] we recorded the identity of the selected body even where both choices earn zero course credit. There was no course credit for this experiment.

Our main contrast compares the probability of choosing the incoming cart in the two conditions in which a participant in the collision speaks. The choice of the incoming cart as the reference fixes the sign of the estimate independently of the allocation. It also allows the comparison to include students who answer that the forces are equal. We registered that contrast with the

three course coordinators before making copies of the response forms. The graph item and the witness comparison were specified as secondary outcomes. Counts of words in explanations were examined after the forms had been read.

III. FILM AND TRANSCRIPT

Two low-friction teaching carts, A and B, each had a mass of 0.48 kg to the displayed precision of the balance. A spring bumper kept their bodies apart. Cart A approached B at approximately 0.35 m/s. Cart B was initially at rest. We use contact to include compression of the bumper, which was attached to A and counted in its mass. The track was horizontal within the resolution of the leveling bubble. We did not infer forces from the subsequent motion. For a common instant in the contact interval, with positive direction along the incoming velocity, the requested relation was

$$F_{A\to B}(t) = -F_{B\to A}(t). \tag{1}$$

Using the force exerted by a body rather than the force on it matters for scoring. A participant can regard A as the aggressor and say that B *receives* the greater force. That answer credits A with exerting the greater force, and is recorded accordingly. We printed the verbs on the response form in full. The terms action and reaction never appeared on the form, since they can suggest a delay [4]. Both quantities in Eq. (1) referred to the same marked frame. The question concerned magnitudes. Opposite signs alone give no basis for choosing a greater magnitude.

Cart C stood on a short second track behind B. Marks on the table and an overhead camera made the transverse separation visible. C was in the picture for the whole approach. At no time was there a rod, string or inclined surface connecting it to either of the other carts. The camera stayed still until the film ended. The floor and the table exerted forces during the recording. The drawing explicitly asked for contacts *between carts*, so neither support belonged on it. A pilot question that asked for forces on carts had elicited reasonable drawings of the supports and was dropped.

Below the last frame was an instrument-style transcript. Each line had a time, a source letter and a payload. Three introductory lines recorded device identification. The last line, when present, appeared 1.20 s after the marked instant and read “I am sorry.” Only the source letter differed among the three versions with a payload. The fourth displayed the same time and a blank source and payload. All four screens stayed up for the same duration. There was no recorded voice. In particular C did not move or light up as its line appeared. The software had assigned the line before the session. The introductory screen said that the log could carry short status messages attributed to each cart.

A criticism at the pilot stage concerned the appropriate recipient. The sentence made no mention of the addressee. An earlier draft named cart B after the apology in all three versions, requiring B in one version to apologize to itself. The final item has no recipient field. Students who read C’s message as an apology on behalf of A were free to retain the pair A and B. We kept such explanations even if they differed from the interpretation that had motivated the study.

IV. PARTICIPANTS, ALLOCATION AND SCORING

We invited students during review sessions in three first-term calculus-based courses. Instruction at the three sites used English for the laboratory component. Each class had covered free-body diagrams and conservation of momentum before its session. A short paper set out the use of anonymized responses and made participation optional. The course instructors stayed outside while forms were completed. Students could complete ordinary revision problems at the same desks. The three teaching-research committees approved the pooled analysis and the immediate explanation of the apparatus. No name, course grade or persistent student identifier was collected.

Of 401 students who attended, 384 consented and received a booklet. Eighty-four, 168 and 132 participants came from the three courses. Each course contributed the same number of students to each of the four conditions within that course, respectively 21, 42 and 33. Booklets went out in permuted blocks of four, in opaque covers, in seat order. We used the same short film on one screen and printed a facsimile of the final screen in each booklet. The projected portion ended just before the transcript. Students worked silently. They could inspect the transcript while answering the comparison. The cover sheet had a colored tab for the proctor, and the tab color was independent of the version inside.

The first response was a choice among A exerts more, B exerts more, and equal magnitudes. On the next page students drew a line for every pair of carts in contact at the marked frame. They then saw two simultaneous force traces from the bumper transducers. Both traces used the same scale in newtons, with one transducer axis reversed so that compression was positive on both displays. Students repeated the magnitude comparison with the traces in view. Booklets had tear-off leaves, and each completed leaf went into a box before the next one opened. A small common serial number linked the three leaves. All 384 returned all three leaves with a usable choice or drawing. Reason statements were op-

tional. The instructor used the last five minutes for an ordinary discussion of the third law and the text generator.

Two readers coded the drawing independently after the source lines had been masked. A line touching C counted as an incident edge. We recorded it even if the drawing also contained the correct edge AB. A label or circle around C with no connecting line did not count. Training used 32 drawings from a separate development class [11]. The readers differed on nine study drawings, seven of them short strokes ending between B and C. A third reader applied the rule in Appendix A without seeing the condition or the first two codes. Our outcome is the adjudicated edge indicator. Its denominator includes a correctly blank area around C. Omitting all drawings without an edge would change the population in the comparison.

For the primary difference in proportions we used the unpooled large-sample standard error. Let p_s be the chance of selecting A when the source is s, with $s = 0$ for the empty line. The estimated difference and its standard error were

$$\widehat{D} = \hat{p}_A - \hat{p}_B, \qquad s_D^2 = \frac{\hat{p}_A(1-\hat{p}_A)}{96} + \frac{\hat{p}_B(1-\hat{p}_B)}{96} \tag{2}$$

Intervals quoted for differences are the estimate plus or minus $1.96 s_D$, expressed in percentage points. Counts in the smaller single-condition cells are available for alternative interval constructions [12, 13]. The two probabilities in Eq. (2) refer to separate students. The graph question was repeated within student and cannot be treated as an independent replication. Because the allocation fractions were identical at the three sites, direct standardization to the pooled site distribution gives precisely the crude difference. The reported standard error ignores the fixed site counts. It describes independent sampling within conditions and does not account for variation between hypothetical new courses.

V. RESULTS

The empty transcript elicited 32 selections of A and ten of B. Thus students who assigned a larger force to one cart preferred the moving one before a source was named. The two changing-source conditions in Table I differ in both incorrect choices. The number selecting A went up by 29 when A apologized. At the same time the number selecting B fell by 26. Three fewer students chose equality. This is a statement about differences between groups. The booklets cannot tell us that any particular student exchanged one incorrect choice for the other.

TABLE I. First comparison at the marked frame. Each row contains 96 students. Entries give the cart exerting the larger force, or equality.

Source of statement	A larger	Equal	B larger
A, incoming	49	42	5
B, struck	20	45	31
C, separate track	33	51	12
Empty	32	54	10

From Eq. (2) the estimated source effect was 30.2 percentage points with an interval from 17.3 to 43.1. The same comparison for the frequency of a correct answer gave a difference of minus 3.1 points (interval minus 17.2 to 11.0). That interval is too wide for an equivalence claim. A test restricted to the subset choosing unequal would make the reversal look larger, but membership in that subset was measured after assignment. We retain the whole group in Fig. 1.

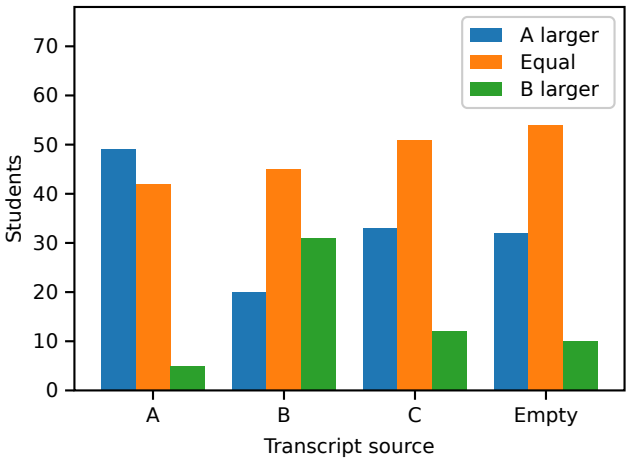


FIG. 1. First comparison by source of apology. The plotted quantities are counts per 96 booklets.

The direction of the primary contrast was the same in each course. In the smallest course, 11 students selected A when it apologized and five when B apologized. The corresponding counts were 21 and eight at the largest site and 17 and seven at the remaining site. The estimates are imprecise individually. An interaction between site and assigned source could be meaningful in a larger set of classes, particularly if the initial collision discussion in one class had used a driver as the grammatical subject [14]. We had three sites and one film, so we did not fit a random slope across institutions.

On the diagram page, naming C increased the occurrence of an edge incident on C by 17.7 points relative to an empty transcript (interval 7.6 to 27.8). The complete distribution in Table II resolves a possible ambiguity. In 18 of the 24 C-source drawings with such an edge, the student retained AB as well. These students added a

contact. Six drew an edge to C and left out AB. Seen from overhead, the gap between the tracks was about three times the width of a cart. The incorrect edge usually passed over that gap.

TABLE II. Contact drawings, classified by presence of AB and any edge to C. The last column is the number choosing equality after seeing force traces.

Source	AB only	AB and C edge	C edge only	No edges	Traces equal
A	83	8	3	2	88
B	84	7	3	2	89
C	70	18	6	2	87
Empty	87	5	2	2	90

The distribution by course for the witness comparison appears in Fig. 2. We show counts against the course-specific denominator to retain the contribution of the smaller course. One student in the smallest empty group has the same weight as one student in the largest C group in the pooled percentage. Reading these as percentages on a common denominator of 96 would give a misleading impression of the absolute frequencies within each site.

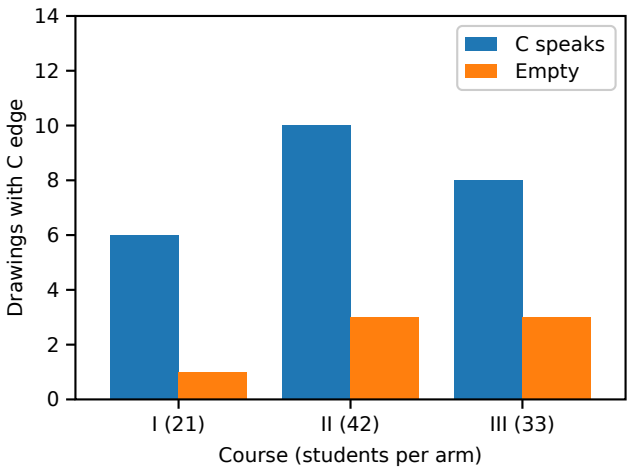


FIG. 2. Contact diagrams carrying an edge to C, by course and source. Each adjacent pair has the same denominator, printed on the horizontal axis.

The graph item broke the connection between attribution and measured magnitude for most respondents. Correct proportions lay between 87/96 and 90/96. The paired changes in Fig. 3 are substantial in each condition, although readout, repetition and order coincide here. We cannot assign the improvement to the traces alone. Rounding each trace to a whole newton changed their relative height in the marked frame by less than one plotted line width. Students could base the correct answer on a visual match without constructing an action-reaction pair. This is one reason we did not use the graph item as the primary endpoint [15, 16].

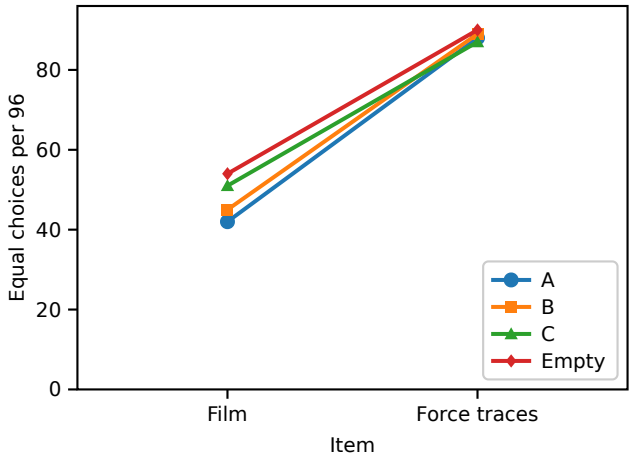


FIG. 3. Students choosing equal forces on the film and graph items. Lines connect repeated observations of the same 96 students per condition; only the group counts are shown.

VI. WHAT THE EXPLANATIONS PERMIT

Of the 384 booklets, 301 contained words in the optional reason box. We transcribed a source letter as a letter even if a respondent had written it inside a speech bubble. Eleven students volunteered that the last line could have referred to something other than the impact. Four made that observation in the silent condition, where the form provided an empty time slot. An empty field evidently did not have a unique reading. It could represent an instrument that had heard nothing, or a report which had yet to be filed. Open-ended work on event lists has encountered the same distinction [17].

A smaller group of drawings put an arrow from C to B and labeled it with the text of the apology. Such a student may have recorded communication on a page restricted to contact, rather than inferred that C crossed the gap. We adjudicated the line as an edge if it met the spatial rule. A force diagram used in a class would normally have arrow labels, but adding a separate penalty here for a verbal label would make the definition of an edge depend on the assigned condition. We therefore interpret the result as an effect on the submitted contact diagram. It has a narrower reach than a claim that students experienced apparent motion of C.

There were also explanations in which the student separated the log from the force comparison in a single sentence. One had A exert the greater force because it approached faster, followed by the observation that B had nothing to apologize for. We retained the first choice as an error. This case illustrates a limitation of a source check asking whether the student noticed the

statement. Noticing it is compatible with treating it as irrelevant, treating it as inaccurate, or using it as the only available evidence. A forced source-recall question between film and diagram would risk turning an unobtrusive manipulation into the assigned task [18, 19].

Examinations present an associated interpretive problem. Suppose two classes receive different narrative versions of an otherwise common third-law question. Equal pass rates do not establish that the versions elicit the same reasoning. One can inspect the relative use of the distractors without trying to infer an individual student’s mechanism from a marginal table. This requires retaining the keyed identity of each distractor, something an undifferentiated mark of zero usually loses. Our estimate applies to a retrospective statement from an object. A passenger speaking from inside a cart would supply a possible percipient witness, altering the evidential status of the same words. Trials with human drivers or persons involved in an actual accident fall outside the population studied here.

VII. SCOPE OF THE INSTRUMENT

Cart labels and directions of travel were fixed. A stronger tendency to read A as active could produce the baseline difference, although it cannot by itself explain a difference between two randomly assigned sender fields. A blank line also offers no estimate of the effect of an equally long, affectively neutral message from each body. A collision report with the payload “record complete” would retain the source cue and remove most of the responsibility cue. Our design assigns a complete apology to a source. It cannot distinguish an effect of authorship of that sentence from an effect of authorship of any short text. Results from speaker-tag tasks suggest that this distinction can matter even in quite sparse displays [20, 21].

Finally, instantaneous equality of the internal forces holds for this apparatus even if the masses are unequal or the collisions inelastic. A question about damage, stopping distance or change of speed can have an unequal answer in physically similar scenes. The condition under which a source may enter the answer is therefore part of the specification of the item. On the present film a transcript can establish which cart has entered a statement into the register. To establish how hard that cart pushed, the frame and the named partner are still required.

APPENDIX A: RESPONSE FORM AND ADJUDICATION DETAILS

The first page said: “At the marked instant the two bumpers are compressed. Compare the magnitude of the horizontal contact force exerted by cart A on cart B with the magnitude of the horizontal contact force exerted by cart B on cart A.” The alternatives repeated both cart letters in each clause. Each copy placed equality in the middle and the incoming-cart choice first. Students could replay the approach in the common projection once, before opening their booklets. There was no facility for scrolling to a different marked frame. One demonstrator kept a paper copy of the projected timing sheet in view to confirm that all four booklet screens began after the final projected frame.

The diagram page showed three squares in the same relative positions as the carts at peak compression, separated by narrow margins to allow a line to terminate at a square. Its instruction was: “Draw a connecting line for each pair of carts in contact at the marked instant. The table and tracks are omitted.” A stroke ending within 2 mm of each of two distinct squares qualified. Crossing a third square did not establish a second edge unless a second stroke terminated there. When the stroke turned within a square and reached a third square, both incident edges qualified. These rules were developed on pilot booklets, where arrows often bent to avoid a printed letter [11]. A connecting stroke with an arrowhead counted irrespective of its direction. The instrument did not ask for a directed graph.

The standard error obtained from the primary pair was 0.06575 on the probability scale. Readers who use Table I to compute a pooled test under equal choice probabilities will obtain a different standard error, appropriate to a different calculation. For the difference on the graph item, one can compare groups in the ordinary way using Table II. For the change *within* a source condition the paired discordance counts are needed. In A, B, C and the empty condition, respectively, the counts moving from an unequal choice to equality were 49, 47, 39 and 39. The counts moving the other way were 3, 3, 3 and 3. These numbers recover both margins of Fig. 3 and do not require independence of successive leaves.

MATERIALS AND ACCESS

The film consists of one approach and a single marked contact frame. The wording, source assignment and decision rules above specify the comparisons available for replication. Institutional consent covered pooled counts and paraphrases. Individual reason statements were kept at the course of origin.

ACKNOWLEDGMENTS

We thank the course technicians who built the second track and the readers who coded the drawings before

the allocation file was opened.

[1] D. Hestenes, M. Wells, and G. Swackhamer, The Physics Teacher **30**, 141 (1992).
[2] R. K. Thornton and D. R. Sokoloff, American Journal of Physics **66**, 338 (1998).
[3] L. Bao, K. Hogg, and D. Zollman, American Journal of Physics **70**, 766 (2002).
[4] L. Bao and J. C. Fritchman, Physical Review Physics Education Research **17**, 020116 (2021).
[5] D. Hammer, American Journal of Physics **64**, 1316 (1996).
[6] E. F. Redish, *Teaching Physics with the Physics Suite* (Wiley, 2003).
[7] M. O. Sabell, Studies in Laboratory Pedagogy **6**, 39 (2017).
[8] M. B. Tarel, Review of Physical Instruction **16**, 105 (2020).
[9] O. I. Grevenet, Studies in Laboratory Pedagogy **7**, 11 (2018).
[10] A. S. Virel, Review of Physical Instruction **15**, 74 (2019).
[11] O. J. Brenet, Laboratory Assessment Notes **9**, 7 (2021).
[12] E. B. Wilson, Journal of the American Statistical Association **22**, 209 (1927).
[13] R. G. Newcombe, Statistics in Medicine **17**, 873 (1998).
[14] S. R. Mornalet, Studies in Laboratory Pedagogy **5**, 92 (2016).
[15] J. M. Planic, Review of Physical Instruction **17**, 24 (2021).
[16] J. L. Trowellin, Laboratory Assessment Notes **6**, 44 (2018).
[17] D. K. Elvane, Laboratory Assessment Notes **10**, 61 (2022).
[18] F. R. Orlunden, Studies in Laboratory Pedagogy **9**, 111 (2020).
[19] T. G. Wexelin, Laboratory Assessment Notes **7**, 83 (2019).
[20] S. A. Rentelac, Review of Physical Instruction **18**, 139 (2022).
[21] I. P. Dovren, Studies in Laboratory Pedagogy **12**, 48 (2023).