*Psicológica (2014), 35*, 479-502.

# An information integration study on the intuitive physics of the Newton's cradle

Nuno Alexandre De Sá Teixeira*, Armando Mónica Oliveira, & Ana Duarte Silva

*University of Coimbra, Portugal*

Newton's cradle, a device consisting of a chain of steel balls suspended in alignment, has been used extensively in physics teaching to demonstrate the principles of conservation of momentum and kinetic energy in elastic collisions. The apparent simplicity of the device allows one to test commonly hold views regarding the intuitive understanding of physics by lay people. We present and discuss the results of two experiments wherein the extension of the chain, the height of release (experiment 1) and the material of the balls (experiment 2) were systematically varied in graphical depictions of the Newton's cradle. Participants had to estimate the height that the last ball in the chain would reach if a collision took place. The outcomes revealed a sophisticated cognitive model wherein the magnitude of the displacement of the target ball increased in direct proportion with the acceleration of the launcher and in inverse proportion with the number of balls in the chain. The results closely mimicked the predicted behavior of a Newton's cradle if the collisions were not perfectly elastic. This isomorphism shows that judgments of physical events are not detached from the environment, as one seldom sees a perfectly elastic collision, and it speaks unfavorably to the hypothesis that, in such tasks, people rely on simple heuristics.

In the 172nd episode of the popular TV show *Mythbusters* (Auty, Dallow, & Williams, 2011), Adam and Jamie set forth to ascertain if it was possible to build a working Newton's cradle using construction wrecking balls. In the course of the show, they built a version of the device composed of five 907kg balls (concrete and steel rebar-filled buoys). Although its efficiency was far from expected, to date this is the largest cradle ever built.

---

* This work was partially supported by Grant SFRH/BPD/84118/2012; Portuguese Foundation for Science and Technology (FCT). Corresponding author: Nuno De Sá Teixeira. E-mail: nuno_desateixeira@fpce.uc.pt

Newton's cradle refers to a device where a series of balls (with the same size and mass; usually steel balls) are suspended so as to be horizontally aligned (see Figure 1, top). When a ball in one end of the chain is held at a certain height and released, so as to collide with the remaining balls, the ball at the opposing end is kicked off at about the same final velocity of the colliding one reaching about the same dropping height. During this event, the intermediate balls are kept sensibly at the same position, apparently unaffected by the collision (see Figure 1, bottom). If more than one ball is held and dropped, the same number of balls leave the chain at the opposing end – the reason why the cradle is also known as *counting device*.

Newton's cradles have been sold as an executive toy and are frequently used in physics classrooms as a demonstration of the conservation of momentum and kinetic energy principles. One common, but mistaken, explanation states that the standard behavior of the cradle is the only outcome that conserves both momentum and kinetic energy, and that would be why the device works as it does. In fact, a series of other theoretical outcomes that also conserve momentum and kinetic energy are conceivable, although seldom observed empirically (cf. Gauld, 2006). Descriptively, the device is the hallmark of the so called *perfectly elastic collisions* – given an encounter between two bodies, an elastic collision is said to take place when the ratio of relative speeds after and before the impact – *coefficient of restitution* –, equals 1. *Elasticity* is here, and in the remaining of this paper, used as a surrogate for the coefficient of restitution, and it varies between 0 – a perfectly plastic collision – and 1 – a perfectly elastic collision (the case portrayed by the Newton's cradle).

The aim of this paper was to determine the cognitive internal model that people use when asked to make predictions regarding the outcome of versions of the Newton's cradle. In doing so, we sought to elucidate the intuitive understanding of collisions held by laypeople and to ascertain if and which grasping of elasticity is implicitly assumed in naïve judgments. Although several studies have focused the sensitivity to variations of elasticity in perceiving dynamic events, it is still to be known if human observers internally model this parameter and, if so, how it is used to assess real world events. Before presenting and discussing the obtained results, we will briefly review previous studies on the perception and understanding of dynamic events.

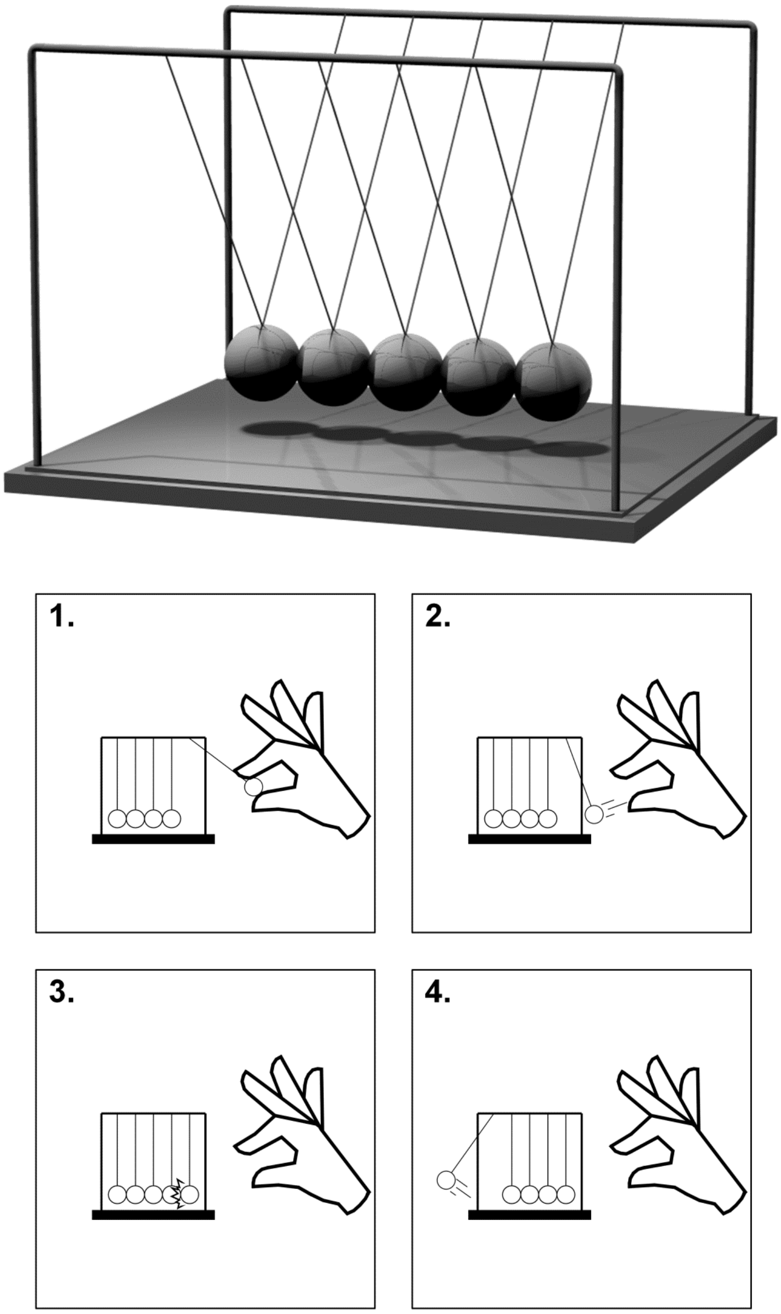


**Figure 1. Top: Example of a Newton's cradle. Bottom: Description of the typical behavior of the device.**

## Perceiving dynamic events

Much of our understanding on how people perceive dynamic events is both historically and conceptually rooted in the pioneer studies lead by Michotte (1963) on the perception of causality. Michotte envisioned and conducted a series of experiments which aimed at showing that *causality* was a directly perceivable quality, in contrast with the views held by Hume (1748/1955). In order to do so, participants were shown simple dynamic events and inquired about their impressions. One of the most studied of such events both by Michotte and following researchers (cf. e.g., Wagemans, Lier, & Scholl, 2006; White, 2009) was coined *launching effect*: two static objects are initially shown some distance apart; at a certain moment, one of the objects (motor object) starts moving towards the other and, upon making contact, it stops its motion while the other object starts moving in the same direction. Given a certain gamut of ideal spatial-temporal conditions, human observers spontaneously report a vivid and irresistible impression that the kinematics of the objects share a causal bound – that is, that the motion of the second object was *caused* by the first one due to a collision or launching action (cf. e.g., Scholl & Tremoulet, 2000; Schlottmann, 2000). Even though Michotte was convinced that causal perception was dissociated from intuitive reasoning of Newtonian mechanics, later researchers came to stress the close coupling between the two (Dittrich & Lea, 1994; Rips, 2011; Runeson, 1983; Weir, 1978; White & Milne, 1997). One of the arguments rose by Michotte for this dissociation came from the outcomes found in his experiments 39 and 40: if the launched object travels at a faster speed than the motor object, the launching impression was lessened instead of strengthened. The stronger impressions of causality were found when the launched object travelled at about 1/4 of the speed of the motor object. As Michotte remarks, "there is no parallelism between the causal impression and the degree of physical force. Instead the impression is rather better and more stable when the efficacy of the 'cause' – as shown in the speed of the projectile – is less!" (Michotte, 1963, p. 109)."

Importantly, as first stressed by Runeson (1983), varying the ratio of velocities of two objects involved in a collision might trigger some other perceptual mechanisms, in the sense that, and under Newtonian physics, it provides information regarding the mass of the objects. Runeson thus suggested that, in accordance with Gibson's ecological theory of perception (Gibson, 1986), the relative ratio of velocities of both objects before and after a collision specified the perception of their relative masses – this prediction stems directly from the Newtonian equations of motion. Moreover, it can be shown that *elasticity* factors out when one rearranges the equations, with the ensuing prediction that elasticity should not influence peoples' perception of relative mass (that is, which object is heavier). This proposal embodies a wider view which came to be known as the *principle of kinematic specification of dynamics* (KSD): when a certain kinematic pattern relates unambiguously to a set of dynamic properties, it is those underlying factors that are perceived – the kinematics are thus said to be an *invariant* (see also Runeson & Frykholm, 1983). This framework fuelled a series of experiments led both by followers of Runeson's ideas (e.g. Runeson & Vedeler, 1993; Runeson, 1995; Runeson et al, 2000) and his critics (e.g., Gilden & Proffit, 1989; Proffitt & Gilden, 1989; Gilden, 1991; Gilden & Proffitt, 1994), which held that simple heuristic reasoning, not KSD, accounted for the perception of relative mass given collision events (see also Hecht, 1996). Regarding the effect of *elasticity*, Todd and Warren (1982) showed that with decreases in the coefficient of restitution, the accuracy of participants to identify the heavier object was proportionally reduced, a result which was taken as a serious rebuttal of the KSD principle (but see Sanford, Mansighka, & Griffiths, 2013). Further evidence that people are sensitive to the implied elasticity in dynamic events was reported by Twardy and Bingham (2002), who showed that naturalness ratings decreased when elasticity was increased beyond its natural bounds (see also, Warren, Kim, & Husney, 1987).

**Overview of the present study**

The main motivation for the experiments to be reported here was to assess the degree to which the sensitivity to elasticity was due to an internal model of the expected behavior of collisions. In order to do so, we relied on an Intuitive Physics task (see, e.g., McCloskey, 1983; McCloskey, Washburn, & Felch, 1983; DiSessa, 1982) where people had to make judgments concerning the expected outcome of versions of the Newton's cradle given a static depiction of its initial conditions (but see Kaiser, Proffitt, & Anderson, 1985, and Kaiser, Proffitt, Whelan, & Hecht, 1992, on the effects of presenting dynamic information on intuitive physics judgments). Furthermore, we made use of the Information Integration Theory and Functional Measurement framework (IIT/FM; Anderson, 1981, 1982), previously shown to be highly effective in revealing algebraic cognitive models in this sort of tasks (see, e.g., Anderson, 1983; Anderson & Wilkening, 1991; Anderson & Schlottmann, 1993; Masin, 2007; Cocco & Masin, 2010; Masin & Rispoli, 2010; Vicovaro, 2012).

The main advantage of this approach is that, given an appropriate physical modelization of the stimuli, specific predictions can be derived which bond the possible interpretation of people's intuitive judgments. In the next paragraphs we develop and discuss such predictions.

**Modeling the physical behavior of the Newton's cradle**

A physical model that fully captures the behavior of the Newton's cradle is still a matter of contention (see Gauld, 2006, for a revision of several proposed models). One formulation, the *independent collision model*, is typically used for virtual animations as it provides a reasonably fair description of the device's observable behavior (Gauld, 2006), albeit at the cost of a simplifying assumption – that the balls in the chain have at least an infinitesimal separation. Despite the shortcomings of this assumption and the model itself for a satisfying explanation in physical terms, it provides an appropriate formulation at the descriptive level. We will thus adopt it as the normative standard in the present paper. Moreover, its mathematical description is relatively straightforward and, being based on the Newtonian equations of motion, epistemologically closer to the existing discussion in the literature regarding the perception and naïve understanding of collisions. Finally, its general logic has an intuitive appeal which, for our level of analysis, makes it appropriate for enquiries regarding the naïve reasoning of humans.

The first step in modeling the device's behavior consists in specifying the final velocity of the striking ball. As the balls in a Newton's cradle behave as pendulums, the velocity at which a ball that is raised and then dropped from a certain height (henceforth referred to as *launcher*) collides with the chain ($V_{Launcher}$) corresponds to its velocity at the equilibrium position, given by the following equation (cf. Alonso & Finn, 1992):

$$(1)\ V_{Launcher} = \sqrt{2 \times g \times l \times (1 - \cos(\alpha))}$$

The parameter $g$ refers to the earth's gravity pull (≈ 980.6 $cm^{-s^2}$), $l$ to the length of the string holding the ball and $\alpha$ to the angle from which it is dropped. The velocities after collision of both the launcher ($U_A$) and the first ball in the chain ($U_B$), given the velocity of the launcher before the collision ($V_A = V_{Launcher}$) can be determined by the following (cf. Alonso & Finn, 1992):

$$(2)\ U_A = \frac{V_A + e \times -V_A}{2}$$

$$(3)\ U_B = \frac{V_A + e \times V_A}{2}$$

In these equations, $e$ refers to elasticity or coefficient of restitution – the ratio of relative speeds after and before a collision – varying between 1, a perfectly elastic collision, and 0, a perfectly plastic collision. Equations 2 and 3 can be used iteratively with the final velocity of the second ball after the first collision, $U_B$, used in place of the initial velocity, $V_A$, as the new "launcher" in the second collision. The same reasoning applies to subsequent collisions eventually reaching the velocity at which the final ball leaves the chain and how high it will travel.

Notice that given a cradle with $K$ balls, there will be $N = K - 1$ collisions. Rearranging equation 3 and accounting for the iteration, it can be shown that the following relation holds:

(4) $U_{Last\ ball} = V_{Launcher} \times \varepsilon^N$

with $\varepsilon = (1 + e)/2$. That is, the velocity at which the final ball leaves the chain ($U_{Last\ ball}$) depends on the velocity at which the first ball strikes it ($V_{Launcher}$) and a parameter dependent on the number and elasticity of collisions. Given a perfectly elastic collision ($e = 1$), then $U_{Last\ ball} = V_{Launcher}$, the standard behavior of the Newton's cradle. Figure 2 depicts the relation expressed in equation 4 for different chain lengths, launcher's drop heights and coefficients of restitution (elasticity).

Several points are worth noticing. When elasticity equals one (the standard behavior of the Newton's cradle), the velocity at which the "launcher" strikes the chain determines alone the velocity at which the last ball exits, irrespective of the number of balls in the chain. In that case $U_{Last\ ball} = V_{Launcher}$. With decreases in elasticity, there is a proportional effect of the number of balls in the chain. The lines corresponding to different chain lengths in figure 2 (left column) can be seen to lie below the line corresponding to $e = 1$, conforming to a linear-fan pattern. Finally, with decreases in the elasticity parameter, a non-linear (quadratic) relation between the velocity of the last ball and chain length emerges (see right column in Figure 2).

Importantly, these trends possess specific statistical markers which, given the inherent variability of human judgments, can be used to assess the degree to which a mental model of this device is isomorphic to the normative model. Thus, given an elasticity parameter less than one, (a) both launcher's velocity and chain's length, taken as factors, should have a main effect and (b) a significant interaction with (c) significant linear-linear and linear-quadratic components.

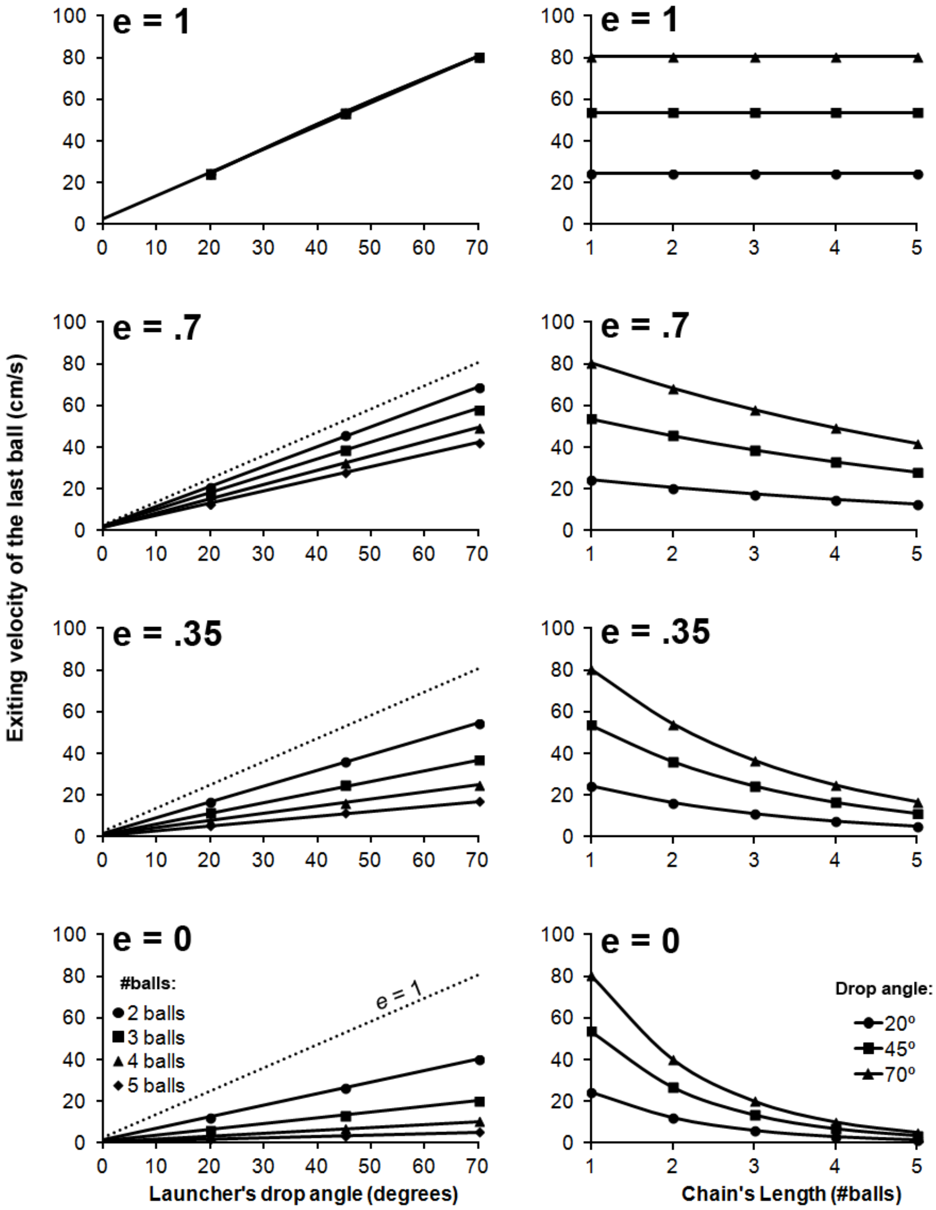


**Figure 2. Last ball's exiting velocity (in cm/s) as a function of launcher's drop height (angle in degrees) and chain's length for elasticity values of 0, 0.35, 0.7 and 1 according to the physical independent collision model.**

## EXPERIMENT 1

### Cognitive algebra of the Newton's cradle

Given the discussion of the adopted normative model, specific scenarios in an intuitive physics task can be outlined. If participants are shown instances of the Newton's cradle with systematic variations of the launcher's drop height and chain's length and asked to indicate how high they would expect the last ball in the chain to reach, the results can be directly confronted with the normative model. A null effect of chain's length in the subjective ratings – *non integration result* – would signal either that people expect perfectly elastic collisions or that they are relying on a simple heuristic considering only launcher's impact velocity. Alternatively, if both launcher's drop height and chain length have an effect on subjective ratings, but no interaction – *parallelism pattern* –, they might possess a mental model with a loose notion of elasticity, albeit distinct from the normative formulation. Finally, if chain's length, launcher's drop height and an interaction between both (with a significant linear-linear component) is found – *linear-fan pattern* –, a case can be made that people do possess a mental model analogue to the normative model and, under certain circumstances, an implicit value of elasticity can be estimated. Moreover, the power operation in equation 4 predicts that a linear-quadratic component of the interaction would tend to be significant for lower values of elasticity. This first experiment aims at establishing which of those scenarios hold.

## METHOD

**Participants.** Eighty psychology students of the University of Coimbra (7 males; 73 females) volunteered for the experiment in exchange for partial course credits. Their ages ranged from 18 to 30 years ($M$ = 20.36; $SD$ = 2.52). All of them had normal or corrected to normal vision and were unaware of the purposes of the task. No participant had formal education in physics besides the obligatory curricula in basic education.

**Stimuli.** Schematic line drawings were used as stimuli. These depicted a side view of a Newton's cradle (see Figure 3). The drawings varied in the number of balls in the chain, from 2 to 5. All balls had a diameter onscreen of 1.6 cm and were shown attached to a string with 5 cm length. In each drawing, one ball, at one or the other end of the chain, was shown raised at a height of 70º, 45º or 20º – *launcher*. At the opposing end of the chain, a curved scale encompassing the allowed trajectory for the last ball was shown with the anchors 0, at the resting position, and 20, at an angle of 90º.

**Procedure and design.** Each participant was shown the entire set of drawings, one at a time and in a random order. For each drawing, the participant had to rate, from 0 (ball at the resting position) to 20 (ball reaching a height of 90º), the height that she/he would expect the last ball to reach if the launcher was dropped from the shown position. They were instructed to imagine, as vividly as possible, a physical device corresponding to the shown depiction and what would happen if the raised ball was dropped. No information was given regarding the material of the balls. The rating was provided by inputting the desired number with the keyboard and confirming each response with the *enter* key. Prior to the experimental task, the participant was allowed to perform a few practice trials encompassing the possible instances of launcher's drop height and chain length. The experiment thus obeyed a full factorial repeated measures design given by 3 (launcher's drop height) × 4 (chain length) × 2 (direction) with 2 repetitions.

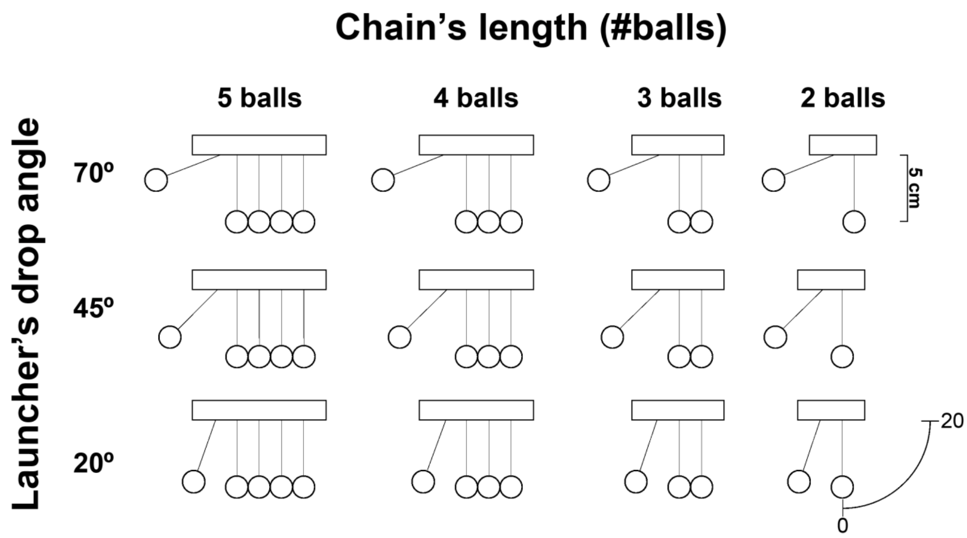


**Figure 3. Stimulus set used in experiment 1 (only the left-to-right implied motions are shown here).**

## RESULTS

The obtained ratings were averaged across replications and direction and subjected to a repeated-measures factorial ANOVA. Whenever the sphericity assumption was not met, the Greenhouse-Geisser correction for the degrees of freedom was used.

Both chain's length, $F(1.35, 106.6) = 91.73$, $p < .001$, $\eta^2_p = .54$, and launcher's drop height, $F(1.22, 96.46) = 227.13$, $p < .001$, $\eta^2_p = .74$, had a significant effect on mean ratings. Ratings increased with increases in drop height and decreased with increasing number of balls in the chain. The effect of launcher's drop height was found to be modulated by chain length, $F(3.9, 308.24) = 40.3$, $p < .001$, $\eta^2_p = .34$. Furthermore, both the linear-linear, $F(1, 79) = 83.93$, $p < .001$, $\eta^2_p = .52$, and the linear-quadratic, $F(1, 79) = 65.06$, $p < .001$, $\eta^2_p = .45$, components of the interaction were found to be significant. These outcomes seem to fulfill all the statistical signs of the normative model and thus suggest that people might have a mental model analogue to the physical behavior of collisions (see Figure 4).

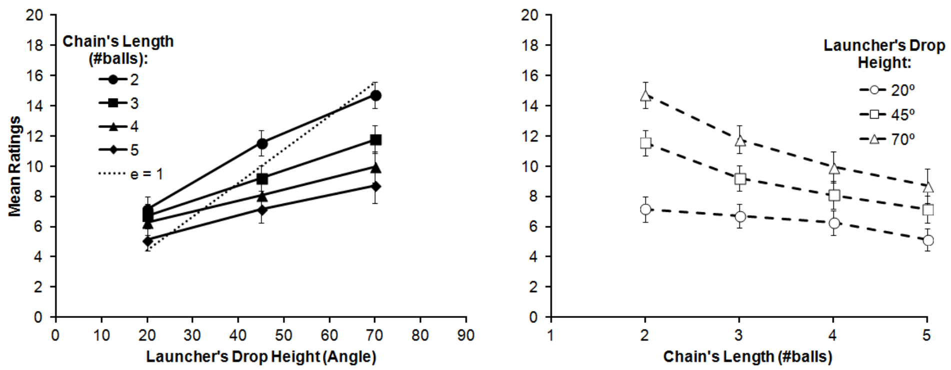


**Figure 4. Mean ratings obtained in experiment 1 as a function of launcher's drop height (angle in degrees; abscissa in left panel, line parameter in right panel) and chain's length (number of balls; abscissa in right panel, line parameter in left panel).**

In equation 4, the last ball's velocity depends on a multiplication between launcher's velocity, proportional to its dropping angle (equation 1) and an elasticity factor dependent on the number of collisions (and henceforth, chain's length). In accordance with the linear-fan theorem (Anderson, 1982), the statistical analysis seems to support a similar algebra in the subjective ratings. The observation of null residuals left behind after fitting a multiplicative model to the empirical data (using Calstat; Weiss, 2006) further supports that a multiplicative pattern seems to be the case. That is, the internal estimate of the height that the target ball would reach had a collision took place ($\rho$) given information about the release height of the launcher ($\varphi$) and the number of balls in the chain ($\psi$), seems to follow the rule given by:

(5) $\rho_{ij} = \varphi_i \psi_j$

Moreover, in accordance with the linear-fan theorem (Anderson, 1982), the response scale seems to be linearly related with the internal estimates:

(6) $R_{ij} = c_0 + c_1 \rho_{ij}$

Although these outcomes suggest that people's judgments qualitatively mimic the expected output of the physical model, there is one important quantitative disparity: namely, the mean ratings are not smaller than the theoretical ratings that would be obtained given a perfectly elastic collision ($e$ = 1). Graphically, that would lead to the line parameters lying below the $e$ = 1 reference line, which is not the case (see Figure 4). Notwithstanding, such a quantitative comparison requires a ratio level measurement, with the lines converging to zero, a condition not warranted with subjective ratings – the linear-fan theorem entails, in general, an interval measure (as implied by equation 6). It is, however, possible to derive ratio measurements given a multiplicative integration rule, as shown algebraically by Masin (2004), by obtaining an empirical estimate of the invariant $c_0$ parameter in equation 6. This quantity, sensitive to the experimental procedure and the way people use the response scale, can then be subtracted from the raw data so as to obtain ratio-level measures of the internal estimate:

(7) $R_{ij} - c_0 = c_1 \rho_{ij}$

This procedure was applied to the collected data, which revealed that $c_0$ amounted to about 5 (that is, participants provided ratings that were on average 5 points above the internal estimate – this has probably to do with a proneness to adjust the answers so as to cover the entire range of the scale).

Figure 5 depicts the ratio level measures (according to equation 7) of participant's responses (data markers) as a function of launcher's final velocity (converted to cm/s according to equation 1; abscissa in left panel and line parameter in right panel) and chain's length (line parameter in left panel and abscissa in right panel). A least-squares procedure was used to fit the normative model to the empirical ratings with elasticity as the only free parameter (*g* and *l* parameters in equation 1 were fixed at 980.6 cm$^{-s^2}$ and 5 cm, respectively). It can be seen that there is a remarkable correspondence between the empirical data and the physical model (lines in figure 5). The elasticity parameter converged to the value of 0.488, with a normalized Root Mean Square Error of .016.

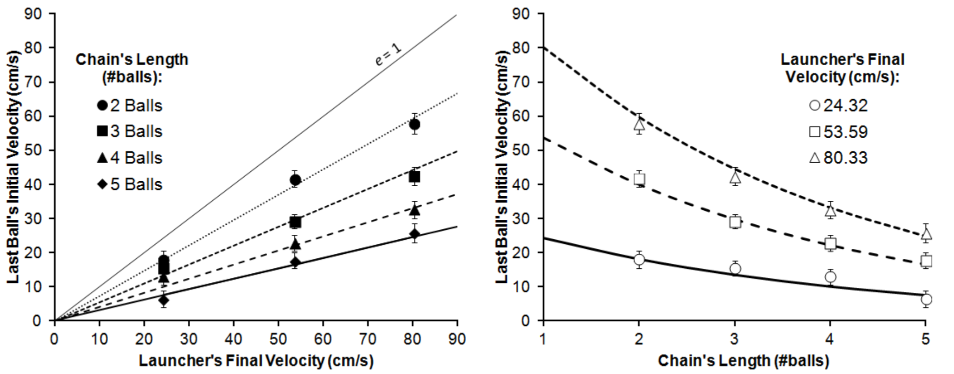


**Figure 5. Ratio measures of participant's responses, converted to cm/s, as a function of launcher's final velocity and chain's length. The lines depict the least-squares fit of the physical model with elasticity as a free parameter (e = 0.488; nRMSE = .016).**

Ratio level ratings were determined (as described in Masin, 2004) on an individual basis. Two participants (2.5%) ignored the variations in chain length and based their responses on the launcher's drop height alone – which might reflect the assumption of a perfectly elastic collision model, or previous experience with the device. Thirteen participants (16%) gave responses that conformed to a parallelism pattern (with launchers dropped from a higher point leading to higher ratings and more balls in the chain reducing the ratings), evidence of an additive cognitive model disparate from the physical formulation. Finally, the remaining participants (81.25%) provided ratings that conformed to the physical model (with e < 1). The individual least squares estimation of elasticity for these participants averaged to 0.3 (*SD* = .26). The difference between participants' ratio ratings and the best least-squares individual solutions (residuals) were subjected to an ANOVA which revealed a null effect of chain's length ($F < 1$), launcher's drop height ($F < 1$) and the interaction term ($F < 1$). It can thus be said that the physical model fully captures the regularities in the participants' responses, supporting an isomorphic cognitive model for collision events.

## DISCUSSION

People's judgments regarding the expected behavior of the Newton's cradle seems to mimic, to a remarkable degree, the physical behavior that would be expected assuming an elasticity value less than 1. Moreover, resting upon the possibility of deriving ratio level measurements of people's judgments, as offered by the IIT/FM framework, it was possible to estimate the mean implicit value of elasticity assumed – about 0.3.

It is important to notice that the Newton's cradle is commonly taken as the hallmark of perfectly elastic collisions, thus with elasticity near 1, used frequently in classrooms as a pedagogic demonstration of some physics principles (but see Gauld, 2006). In their intuitive judgments, people seem to expect a much lower elasticity value. Importantly, this seems to be a default value assumed by observers, as the depictions were devoid of any information that could be used to infer the coefficient of restitution. On the one hand, being schematic depictions, the stimuli provided no information about the balls' materials. On the other hand, no collision was ever actually observed – the judgments relied alone on peoples' expectations. In the world, with no other information, elasticity can only be ascertained *a posteriori*, based upon the ratio of relative velocities of the objects before and after a collision. Our data thus suggests that an internal mental model of collisions instantiates such events with an implicit elasticity anchored at a default value of about 0.3.

Use of a curved scale (which was meant to help participants tracking the allowed trajectory of the target ball) is at odds with common practice and might be suspected of inducing a non-linearity in the response. This possibility is fairly ruled out not just by strict compliance of the observed patterns with the linear-fan theorem (both graphically and statistically), as by the fact that an important percentage of participants (18.75%) actually deviated from the normative model to provide parallel patterns – parallelism having a core role in the validation of response scales as linear (Anderson, 1982, p. 15).

That some participants seemed to follow a disparate model might signal the existence of individual differences in the formulation of the mental models or the employment of different strategies to comply with the task. Be the case as it may, this outcome suggests that the activation of a mental model analogue to the normative formulation might depend on some other factors.

## EXPERIMENT 2

### Cognitive algebra of Newton's cradle with different materials

Experiment 1 provided evidence that humans possess a mental model of collision events isomorphic to a physical formulation. This conclusion was supported both qualitatively (by inspecting the graphical patterns and statistical signs) and quantitatively (by successfully fitting the physical model to the collected responses). Moreover, it was possible to estimate an implicit elasticity value, which averaged at about 0.3. Given these outcomes, one can ask if this implicit elasticity parameter is sensitive to the same factors that modulate the physical coefficient of restitution. One such factor is the materials of the objects involved in the collision. It is important to notice at this point that the coefficient of restitution (or elasticity, as used throughout this paper) is a property of collisions, not objects. This means that, although objects with different materials result in different elasticity values (everything else being equal), a specific coefficient of restitution cannot be unambiguously given to a specific material. For instance, the shape and geometry of objects made with the same material influences the elasticity of the collision (consider, to provide a widespread example, the different outcomes that would be obtained with a Newton's cradle made with steel balls or with steel bells). In this experiment we sought to determine the degree to which the cognitive model of collisions was changed when shown a Newton's cradle with balls made of steel, wood and rock. These materials were chosen because, overall, they would result in different behaviors determined by fairly distinct elasticity values (the software *Interactive Physics* models the behavior of these materials by assuming elasticity values of 0.95, 0.5 and 0.2 to steel, wood and rock, respectively, although these values should be taken as a heuristic generalization). Moreover, these specific materials are relatively straightforward to represent pictorially, by simply varying the texture of the balls. We hypothesized that people, when shown these variations of the Newton's cradle, would adjust their ratings to conform to different implicit values of elasticity.

## METHOD

**Participants.** 33 students (32 females; 1 male) of the University of Coimbra volunteered for the experiment in exchange for partial course credits. Their ages ranged from 18 to 30 years ($M$ = 19.48; $SD$ = 3.09). All participants had normal or corrected to normal vision and were unaware of the purposes of the task. No participant had formal education in physics besides the basic obligatory curricula. Furthermore, no participant declared being familiar with the Newton's cradle.

**Stimuli.** Pictorial colored representations of a three-dimensional realistic Newton's cradle were used as stimuli. Because merely manipulating the textures of the spheres in the previous setting was found to induce ambiguities in stimuli's interpretation (with naïve observers reporting that textured balls resembled planets in a toy planetarium) the choice was made to use overall realistic depictions instead of schematic drawings. All pictures were created and rendered with the *Poser* software (see Figure 6 for examples). The depictions showed the device with 2 to 5 balls in the chain and either the leftmost or rightmost ball raised to a height of 20º, 45º or 70º. The length onscreen of the strings holding the balls was about 8 cm. The balls were textured to suggest either steel, wood or rock balls.

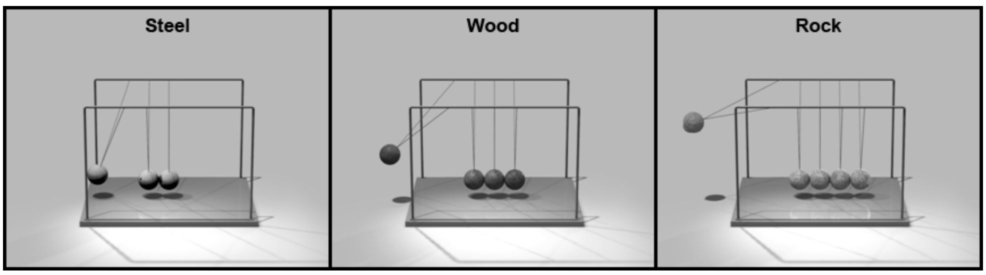


**Figure 6. Grayscale examples of the stimuli used in experiment 2.**

**Procedure and design.** The procedure was in all respects alike the one employed in experiment 1 with the following exceptions. A linear visual analogue scale, shown at the bottom of the screen, was used instead of the curved scale used in experiment 1. The scale was anchored at the extreme values with a pictogram of a ball at rest, at the left end, and a ball rose to a height of 90º, at the right end. Participants were instructed to provide the ratings by clicking with the mouse cursor on the location of the scale that corresponded to the imagined height that the last ball in the chain would reach had the launcher been dropped, taking as reference the positions depicted at the extreme values. Each participant repeated the task three times, one for each depicted material. The order at which each material was shown was counterbalanced by a Latin squares design. The experiment thus obeyed a full factorial repeated measures design given by 3 (material [blocked]) × 3 (Launcher's drop height) × 4 (chain's length) × 2 (direction of implied motion) with each trial being presented twice per participant.

## RESULTS

The obtained ratings were averaged across replications and direction and subjected to a full factorial repeated measures ANOVA. Main effects of material, $F(2, 64) = 15.57$, $p < .001$, $\eta^2_p = .33$, launcher's drop height, $F(1.2, 38.67) = 133.79$, $p < .001$, $\eta^2_p = .8$, and chain's length, $F(1.37, 44.1) = 144.63$, $p < .001$, $\eta^2_p = .82$, were found. A significant interaction between launcher's drop height and chain's length was also found, $F(3.04, 97.18) = 19.88$, $p < .001$, $\eta^2_p = .38$, with significant linear-linear, $F(1, 32) = 36.3$, $p < .001$, $\eta^2_p = .53$, and linear-quadratic components, $F(1, 32) = 28.6$, $p < .001$, $\eta^2_p = .47$. Finally, material was found to interact with launcher's drop height, $F(2.23, 71.37) = 6.99$, $p = .001$, $\eta^2_p = .18$. These outcomes suggest that the same overall pattern found in experiment 1, closely mimicking the physical model, was replicated here. To further ascertain if such was the case for each material condition, separate ANOVAs were performed over the ratings obtained in the steel, wood and rock instances. Launcher's drop height had a significant effect in the steel, $F(1.18, 37.76) = 78.85$, $p < .001$, $\eta^2_p = .71$, wood, $F(1.29, 41.56) = 113.39$, $p < .001$, $\eta^2_p = .78$, and rock conditions, $F(1.18, 37.72) = 91.98$, $p < .001$, $\eta^2_p = .74$. Likewise, chain's length had an effect in all the material conditions: Steel, $F(1.82, 58.27) = 96.07$, $p < .001$, $\eta^2_p = .75$; Wood, $F(1.49, 47.73) = 81.29$, $p < .001$, $\eta^2_p = .72$; Rock, $F(1.57, 50.27) = 96.17$, $p < .001$, $\eta^2_p = .75$. Finally, for all materials, the interaction term and both the linear-linear and linear-quadratic components were also significant: Steel, $F(3.88, 124.23) = 10.6$, $p < .001$, $\eta^2_p = .25$, bilinear, $F(1, 32) = 30.32$, $p < .001$, $\eta^2_p = .49$, linear-quadratic, $F(1, 32) = 8.26$, $p < .001$, $\eta^2_p = .2$; Wood, $F(4.15, 137.78) = 13.95$, $p < .001$, $\eta^2_p = .3$, bilinear, $F(1, 32) = 32.57$, $p < .001$, $\eta^2_p = .5$, linear-quadratic, $F(1, 32) = 20.24$, $p < .001$, $\eta^2_p = .39$; Rock, $F(3.15, 100.89) = 7.76$, $p < .001$, $\eta^2_p = .2$, bilinear, $F(1, 32) = 14.37$, $p = .001$, $\eta^2_p = .31$, linear-quadratic, $F(1, 32) = 17.61$, $p < .001$, $\eta^2_p = .35$.

As in experiment 1, the ratings obtained for each material condition were analyzed with *Calstat* (Weiss, 2006) to test the fit of a multiplicative model (as entailed by equation 4), which revealed that for all materials the residuals were not significant ($F < 1$ in all conditions).

Taken together, these results once again support an isomorphism between the observers' cognitive model and the physical behavior. As a multiplicative model analogue to the one expressed in equation 4 is supported, it is legitimate to derive ratio measures from the raw ratings (equations 5-7). Figure 7 depicts the ratio ratings obtained at the group level (data markers) for the steel (panels A and B), wood (panels C and D) and rock (panels E and F) conditions together with the best least-squares fit of the physical model. It can be seen that the physical model only roughly captures the structure of the empirical data. Even so, the best least-squares solutions converged to estimates of elasticity of 0.26 for steel (normalized RMSE = .03), 0.3 for wood (normalized RMSE = .04) and 0.03 for rock (normalized RMSE = .03).

The procedure used to estimate the implicit elasticity value for the group data was repeated for each participant. In all conditions some participants provided responses that conformed to a parallelism pattern, thus deviating from the physical model. Importantly, these trends were neither related to specific individuals nor with the specific orders of presentation of the different materials (except for one participant who systematically responded according to an additive model). That is, any participant was as likely to provide responses that conformed to a parallelism pattern in any condition and irrespective of it being the first, middle or last block he/she was subjected to. Therefore, 82% of the participants in the steel condition, 85% in the wood condition and 85% in the rock condition responded as if closely following a model analogue to the physical formulation. For these participants, the individual best least-squares solution was found and the residuals subjected to repeated measures ANOVAs. Significant residuals from the launcher's drop height were found in the steel, $F(1.43, 42.9) = 7.64$, $p = .004$, and wood conditions, $F(2, 54) = 3.16$, $p = .05$, but not in the rock condition, $F(2, 54) = 1.81$, $p = .18$. The residuals from the chain's

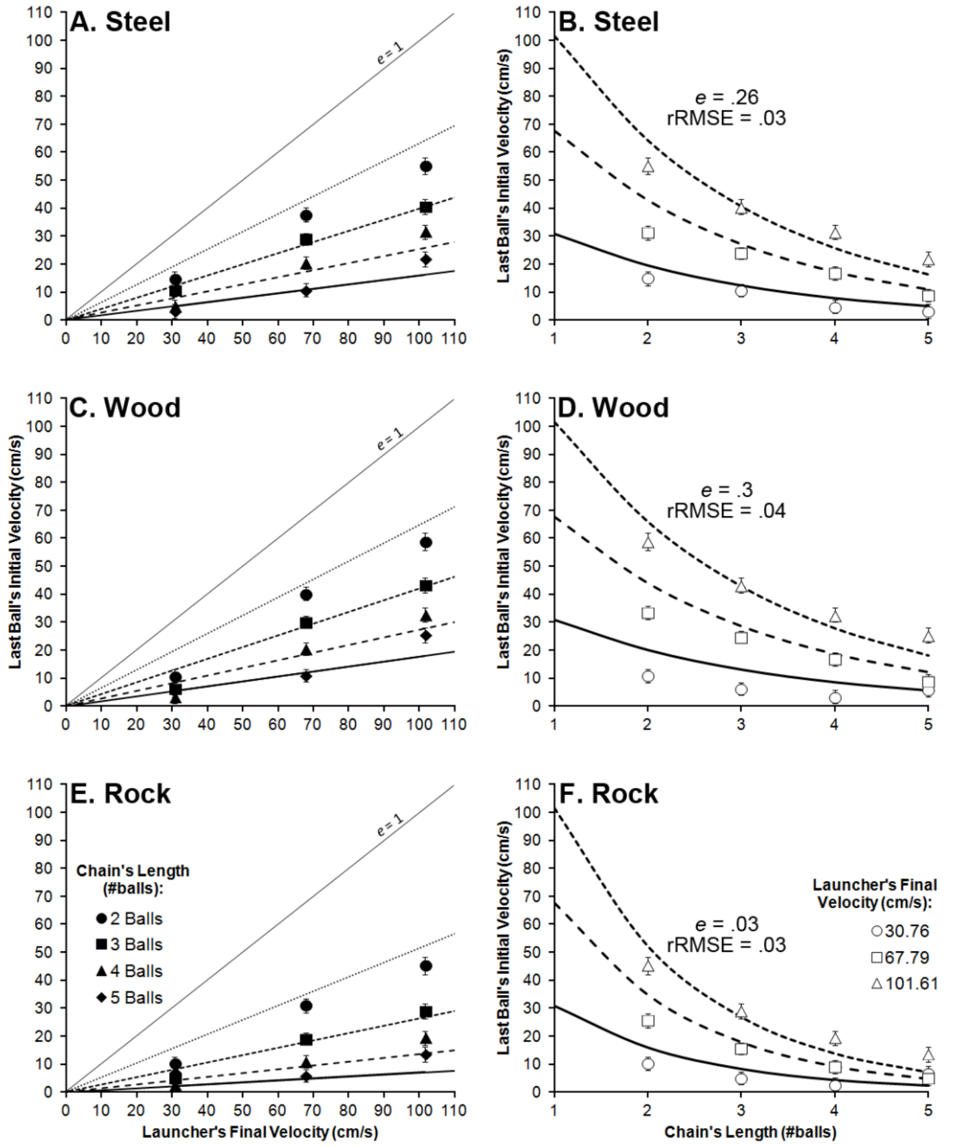


**Figure 7. Mean ratio measures derived from the ratings in experiment 2 (data markers) as a function of launcher's drop height (abscissa in the left panels; line parameter in the right panels) and chain's length (abscissa in the right panels; line parameter in the left panels) for the steel (panels A and B), wood (panels C and D) and rock (panels E and F) conditions. The lines depict the best least-squares solution for each material.**

length were not significant in all conditions: Steel, $F(1.39, 41.95) = 1.69$, $p = .2$; Wood, $F(1.4, 37.88) < 1$; Rock, $F(1.56, 42.16) = 1.77$, $p = .19$. There were no interactions in the residuals. Overall, the results of the residuals test suggest that the physical model fails to account for the effect of the launcher's drop height and only when steel or wood balls are shown. Importantly, the effect of chain's length, directly linked with the implicit elasticity, seems to be fully captured by the physical model, which makes legitimate that an estimate of its value is derived. The mean elasticity value was found to be about 0.33 ($SD = .35$) for the steel, 0.39 ($SD = .34$) for the wood and 0.14 ($SD = .42$) for the rock condition. Moreover, these values were found to result in a significant difference, $F(1.59, 50.94) = 5$, $p =. 016$, $\eta^2_p = .14$, mainly due to the rock's elasticity which was found to be significantly lower than both the steel's and wood's estimate.

## DISCUSSION

The outcomes of experiment 2 replicated to a certain extent the results found in experiment 1. In general, people's judgments were remarkably isomorphic to Newtonian mechanics. Notwithstanding, the fits of the physical model were not as good as in experiment 1. This seems to be mostly due to the effect of launcher's impact velocity, which was not fully captured by equation 1, and not to the implicitly assumed elasticity value, as instantiated in equation 4. It might be that the pictorial three-dimensional cues present in the stimuli of experiment 2 lead to some difficulties in interpreting unambiguously the implied size and/or distance of the device (thus impacting on the $l$ parameter in equation 1). This issue could be clarified by presenting real devices to the participants.

Regarding the estimated values of elasticity, the results of experiment 2 are somewhat mixed. On the one hand, varying the material of the balls lead to some differences in the implicit elasticity in the expected direction – namely, *rock* balls resulted in a significantly lower elasticity. On the other hand, the found elasticity values were not as widespread as expected and clustered around a value of about 0.28. These outcomes strength the hypothesis that a value of about 0.3 might be a strong anchor in regards to the expected elasticity of collisions.

Lastly, a surprisingly stable percentage of participants (about 15%) made judgments that deviated from the Newtonian mechanics. Importantly, these deviations did not seem to reflect systematic individual differences, as the same participant was as likely to provide judgments conforming to a parallelism pattern or to the normative linear-fan pattern. Also, these deviations do not seem to depend on prior experience with the task, as parallelism was equally likely to occur with any material, irrespective of the presentation order. Instead, these outcomes suggest that participants often resorted to varying strategies in construing their representation of the stimuli and the task (for a similar argument see, e.g., Schlottmann & Wilkening, 2011). Interestingly, that being the case, the ensuing conclusion is that even though humans do possess a mental model analogue to Newtonian mechanics, they do not necessarily use it and can favor different response strategies, depending on the circumstances. In fact, there is some evidence (see, e.g., Anderson, 1996) that an additive model, as implied by the parallelism pattern, might be ontologically simpler and, under some circumstances, good enough and easier to employ for some judgments.

## GENERAL DISCUSSION AND CONCLUSION

This paper presented the results of two experiments on the intuitive physics of collision events, using depictions of the Newton's cradle. Participants had to estimate how high the last ball in the chain would reach given variations in the launcher's drop height and the number of balls. The results were compatible with a mental model which strikingly mimics the outcomes that would be expected with a real device given an elasticity value of less than one. By taking advantage of the logic of IIT/FM, it was furthermore possible to estimate the elasticity parameter implicit in people's judgments. This value was found to be about 0.3. Although it was shown to be sensitive to variations in the implied material of the balls, the internal elasticity inputted by the participants was shown to be anchored around that value. From an ecological point of view, an elasticity value within this range could be expected, since perfectly elastic collisions are seldom observed and near perfectly plastic collisions are much more frequent. A default value of 0.3 could thus have been internalized in order to reflect knowledge on the natural statistics of our ecological environment. Lower elasticity values in our environment are thus more likely and our internal estimates seem to reflect knowledge of that probability distribution. These ideas are in line with the arguments raised by Sanford, Mansinghka and Griffiths (2013).

As a final remark, our findings explain much of the appeal of Newton's cradle, either seen as a toy or a classroom demonstration: the near perfectly elastic behavior of the device would certainly be, as it is, awe-inspiring to a being who expects, based on a prior estimation, a low elasticity collision. That the intriguing behavior of the Newton's cradle reflects but a misestimation of one single parameter rather than an inappropriate cognitive model has profound implications to conceptions of the layman physics and perspectives on the teaching of physics.

## REFERENCES


Alonso, M., & Finn, E. J. (1992). *Physics*. Harlow: Pearson – Prentice Hall.

Anderson, N. H. (1981). *Foundations of Information Integration Theory*. Boston: Academic Press.

Anderson, N. H. (1982). *Methods of Information Integration Theory*. Boston: Academic Press.

Anderson, N. H. (1983). Cognitive algebra in intuitive physics. *Advances in Psychology, 11*, 229-253.

Anderson, N. H. (1996). *A Functional Theory of Cognition*. Hillsdale, N.J: L. Erlbaum Associates.

Anderson, N. H., & Wilkening, F. (1991). Adaptive thinking in intuitive physics. In N.H. Anderson (Ed.), *Contributions to information integration theory. Vol 3: Developmental* (pp. 1-42). Hillsdale, NJ: Erlbaum.

Auty, J., Dallow, A., & Williams, L. (2011). Wrecking ball baloney [Television series episode]. In J. Hyneman & A. Savage, *Mythbusters*. Silver Spring: Discovery Channel.

Cocco, A., & Masin, S. C. (2010). The law of elasticity. *Psicológica, 31*, 647-657.

DiSessa, A. A. (1982). Unlearning Aristotelian physics: A study of knowledge-based learning. *Cognitive Science, 6*, 37-75.

Dittrich, W. E., & Lea, S. E. (1994). Visual perception of intentional motion. *Perception, 23*, 253-268.

Gauld, C. F. (2006). Newton's cradle in physics education. *Science & Education, 15*, 597-617.

Gibson, J. J. (1986). *The Ecological Approach to Visual Perception*. New Jersey: Lawrence Erlbaum Associates.

Gilden, D. L. (1991). On the origins of dynamical awareness. *Psychological Review, 98*, 554-568.

Gilden, D. L., & Proffit, D. R. (1989). Understanding collision dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 372-283.

Gilden, D. L., & Proffitt, D. R. (1994). Heuristic judgements of mass ratio in two-body collisions. *Perception & Psychophysics, 56*, 708-720.

Hecht, H. (1996). Heuristic and invariants in dynamic event perception: Immunized concepts of nonstatements? *Psychonomic Bulletin & Review, 3*, 61-70.

Hume, D. (1955). *An inquiry concerning human understanding*. Indianopolis: Bobbs-Merrill. (Original work published in 1748).

Kaiser, M. K., Proffitt, D. R., & Anderson, K. (1985). Judgments of natural and anomalous trajectories in the presence and absence of motion. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 795-803.

Kaiser, M. K., Proffitt, D. R., Whelan, S. M., & Hecht, H. (1992). Influence of animation on dynamical judgments. *Journal of Experimental Psychology: Human Perception and Performance, 18*, 669-689.

Masin, S. C. (2004). Tests of functional measurement theory for multiplicative models. In A. M. Oliveira, M. Teixeira, G. F. Borges, & M. J. Ferro (Eds.), *Proceedings of the* *twentieth annual meeting of the International Society for Psychophysics*. Coimbra: International Society for Psychophysics (pp. 447-452).

Masin, S. C. (2007). Functional measurement in physics. *Teorie & Modelli, 12*, 277-280.

Masin, S. C., & Rispoli, S. (2010). The intuitive law of buoyancy. In A. Bastianelli & G. Vidotto (Eds.), *Proceedings of the twenty-sixth annual meeting of the International Society for Psychophysics*. Padua, Italy: International Society for Psychophysics (pp. 315-320).

McCloskey, M. (1983). Intuitive physics. *Scientific American, 284*, 114-123.

McCloskey, M., Washburn, A., & Felch, L. (1983). Intuitive physics: The straight-down belief and its origin. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 9*, 636-649.

Michotte, A. (1963). *The Perception of Causality*. New York: Basic Books.

Proffitt, D. R., & Gilden, D. (1989). Understanding natural dynamics. *Journal of Experimental Psychology: Human Perception and Performance, 15*, 384-393.

Rips, L. J. (2011). Causation from perception. *Perspectives on Psychological Science, 6*, 77-97.

Runeson, S. (1983). On visual perception of dynamic events. *Acta Universitatis Upsaliensis: Studia Psychologica Upsaliensia* (Serial No. 9). Stockholm, Sweden: Almqvist & Wicksell.

Runeson, S. (1995). Support for the cue-heuristic model is based on suboptimal observer performance: Response to Gilden and Proffitt (1994). *Perception & Psychophysics, 57*, 1262-1273.

Runeson, S., & Fryckholm, 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*, 617-632.

Runeson, S., & Vedeler, D. (1993). The indispensability of precollision kinematics in the visual perception of relative mass. *Perception & Psychophysics, 53*, 617-632.

Runeson, S., Juslin, P., & Olsson, H. (2000). Visual perception of dynamic properties: Cue heuristics versus direct-perceptual competence. *Psychological Review, 107*, 525-555.

Sanford, A. N., Mansighka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and newtonian mechanics for colliding objects. *Psychological Review, 120*, 411-437.

Schlottmann, A. (2000). Is perception of causality modular? *Trends in Cognitive Sciences, 4*, 441-442.

Schlottmann, A., & Wilkening, F. (2011). Judgment and decision making in young children. In M. Dhami, A. Schlottmann, M. Waldmann (Eds.), *Judgment and decision making as a skill* (pp. 55-83). Cambridge University Press.

Schlottmann, A., & Anderson, N. H. (1993). An information integration approach to phenomenal causality. *Memory and Cognition, 21*, 785-801.

Schlottmann, A., Harman, R. M., & Paine, J. (2012). Averaging and adding in children's worth judgments. *Psicológica, 33*, 495 – 513.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*, 299-309.

Todd, J. Y., & Warren, W. H., Jr. (1982). Visual perception of relative mass in dynamic events. *Perception, 11*, 325-335.

Twardy, C. R., & Bingham, G. P. (2002). Causation, causal perception, and conservation laws. *Perception & Psychophysics, 64*, 956-968.

Vicovaro, M. (2012). Intuitive physics of collision effects on simulated spheres differing in size, velocity, and material. *Psicológica, 33*, 451-471.

Wagemans, J., Lier, R. V., & Scholl, B. J. (2006). Introduction of Michotte's heritage in perception and cognition research. *Acta Psychologica, 123*, 1-19.

Warren, W. H., Jr., Kim, E. E., & Husney, R. (1987). The way the ball bounces: Visual and auditory perception of elasticity and control of the bounce pass. *Perception, 16*, 309-336.

Weir, S. (1978). The perception of motion: Michotte revisited. *Perception, 7*, 247-260.

Weiss, D. J. (2006). *Analysis of variance and functional measurement: A practical guide*. NY: Oxford University Press.

White, P. A. (2009). Perception of forces exerted by objects in collision events. *Psychological Review, 116*, 580-601.

White, P. A., & Milne, A. (1997). Phenomenal causality: Impressions of pulling in the visual perception of objects in motion. *American Journal of Psychology, 110*, 573-602.

(Manuscript received: 10 November 2013; accepted: 8 January 2014)