Acta Psychologica 186 (2018) 133–144

Contents lists available at ScienceDirect

Acta Psychologica

journal homepage: www.elsevier.com/locate/actpsy

Causal reports: Context-dependent contributions of intuitive physics and visual impressions of launching

Michele Vicovaro*

Department of General Psychology, University of Padua, Italy

ARTICLE INFO

Keywords:
Causal perception
Causal inference
Intuitive physics
Impetus
Launching effect
Michotte

ABSTRACT

Everyday causal reports appear to be based on a blend of perceptual and cognitive processes. Causality can sometimes be perceived automatically through low-level visual processing of stimuli, but it can also be inferred on the basis of an intuitive understanding of the physical mechanism that underlies an observable event. We investigated how visual impressions of launching and the intuitive physics of collisions contribute to the formation of explicit causal responses. In Experiment 1, participants observed collisions between realistic objects differing in apparent material and hence implied mass, whereas in Experiment 2, participants observed collisions between abstract, non-material objects. The results of Experiment 1 showed that ratings of causality were mainly driven by the intuitive physics of collisions, whereas the results of Experiment 2 provide some support to the hypothesis that ratings of causality were mainly driven by visual impressions of launching. These results suggest that stimulus factors and experimental design factors − such as the realism of the stimuli and the variation in the implied mass of the colliding objects − may determine the relative contributions of perceptual and post-perceptual cognitive processes to explicit causal responses. A revised version of the impetus transmission heuristic provides a satisfactory explanation for these results, whereas the hypothesis that causal responses and intuitive physics are based on the internalization of physical laws does not.

1. Introduction

In one of Michotte's (1963) seminal experiments on the perception of causality, observers were presented with two small, horizontally aligned squares; at a point in time one square (*A*) started moving towards the other (*B*). When *A* made contact with *B*, *B* started moving with the same velocity as *A*, whilst *A* came to a halt (see Fig. 1). The vast majority of observers described this scene by saying that *A* “launched” or “kicked” *B*—that is, that the motion of *A* had *caused* the motion of *B*. This phenomenon was called the *launching effect*. Through a series of ingenious experimental demonstrations, Michotte (1963) showed that the launching effect is a genuinely visual phenomenon, because the necessary and sufficient conditions for its occurrence all relate to the perceptual properties of the scene. That is, the effect occurs when the *perceived* scene satisfies certain requirements; for example, two distinct objects must be present and their motions must exhibit perceptual continuity. In contrast, Michotte showed that the launching phenomenon was not related to either observer's *knowledge* of collisions or the degree of consistency between the simulated collisions and the physical laws of collisions: observers reported visual impressions of launching even when relationships between the physical motions of *A* and *B* were inconsistent with physical laws of collisions. More recently, Michotte's Gestalt-theoretic account of the perception of causality has been reinterpreted in terms of *modularity*. In other words, the launching effect is conceived of as the result of a visual module which is impervious to learning, past experience, and high-level cognitive processes (Leslie & Keeble, 1987; Scholl & Tremoulet, 2000; cf. Rips, 2011).

Michotte's ideas contrast markedly with the empiricist approach to causal relations. The empiricist philosopher David Hume argued (Hume, 1977) that causality cannot be directly perceived and that subjective impressions of causality stem from acquired knowledge about the relationship between separate motions that are characterized by spatiotemporal contiguity. Shortly after the publication of Michotte's work, the empiricist account of perceived causality was supported by research emphasizing the roles of learning (Brown & Miles, 1969; Gruber, Fink, & Damm, 1957; Powesland, 1959) and individual differences (Beasley, 1968; Boyle, 1960; Gemelli & Cappellini, 1959) in the perception of causality. Although the value of Michotte's studies is widely acknowledged in contemporary vision research (see Wagemans, van Lier, & Scholl, 2006), there is as yet no consensus on the relative contributions of low-level visual processes and learning and past

---

* Università di Padova, Dipartimento di Psicologia Generale, via Venezia 8, I-35131 Padova, Italy.
E-mail address: michele.vicovaro@unipd.it.

https://doi.org/10.1016/j.actpsy.2018.04.015
Received 29 September 2017; Received in revised form 13 April 2018; Accepted 23 April 2018

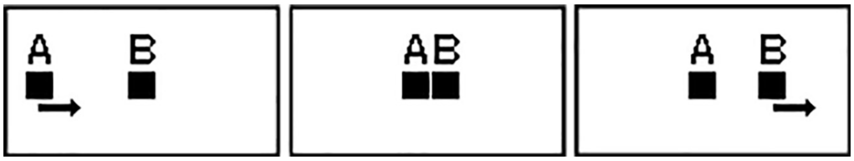

Fig. 1. Three frames of a Michottean collision. The letters *A* and *B* and arrows were added to indicate which objects are moving in the three stages of the collision event.

experience to perceived causality (see Hubbard, 2013a, 2013b). This lack of consensus is partly due to a methodological problem. Most research in this domain has been conducted using explicit measures, for example by asking participants to rate the extent to which the motion of *B* appears to be *caused* by the motion of *A*. However, in the words of Choi and Scholl (2006, p. 93) “explicit reports are always sensitive in principle to extra-perceptual factors, and one of the most serious concerns is that verbal reports reflect not only what subjects are *seeing* but also their higher-level interpretations and judgments.” In other words, learning and past experience might be relevant to explicit reports of causality – and hence to the explicit post-perceptual processing of the stimuli – rather than to perceived causality per se.

In order to bypass the seemingly inescapable problem of the contribution of explicit post-perceptual processing to *explicit* reports of causality, researchers have recently focused on *implicit* measures of causal perception. For instance, it has been shown that perception of the launching effect implies a distortion of the perceived distance between *A* and *B* (Buehner & Humphreys, 2010; Scholl & Nakayama, 2004) and a modification of the perceived trajectory of the apparent motion of *A* (Kim, Feldman, & Singh, 2013). Moors, Wagemans, and de-Wit (2017) showed that visual stimuli that normally elicit a launching effect enter awareness faster than similar stimuli that do not elicit a causal impression. Rolfs, Dambacher, and Cavanagh (2013) showed that the launching effect was subject to specific retinotopic visual adaptation. These studies highlighted behavioral consequences of the visual perception of launching that emerge despite the lack of explicit reference to causality in the experimental instructions, and so they suggest that causal perception stems from automatic, low-level visual processing of Michottean collisions (i.e., collisions like that depicted in Fig. 1).[1] Further support for this claim has been provided by neurophysiological and neuroimaging studies showing that distinct brain regions are involved in causal perception and causal reasoning (Fonlupt, 2003; Roser, Fugelsang, Dunbar, Corballis, & Gazzaniga, 2005; cf. Straube & Chatterjee, 2010).

1.1. A connection between causal reports and impetus transmission?

Implicit measures allow researchers to explore genuine visual impressions of causality independently from the influence of explicit post-perceptual causal reasoning. Nevertheless, everyday causal reports appear to be based on a blend of perceptual and cognitive processes (Schlottmann, 2000, 2001; Schlottmann & Anderson, 1993). Our explicit responses about the possible existence of a causal relationship between two events are likely to be driven not only by our immediate visual impressions, but also by what we *know* about the events. According to *physicalist models* of causal cognition, causal inferences are based on analogies with the physical world; in other words, events are believed to be causally related if there is a plausible physical *mechanism* to explain such a relationship (Bullock, Gelman, & Baillargeon, 1982; Schlottmann, 1999; Schultz, Fisher, Pratt, & Rulf, 1986; Wolff, 2007). In this context, the term “mechanism” refers to people's intuitive understanding of a physical event, which may not correspond with the relevant physical laws (e.g., diSessa, 1993; McCloskey, 1983). The *property transmission* hypothesis (White, 2009a) reflects a physicalist account of causal cognition: causal inferences are drawn when some property is implicitly or explicitly believed to be transmitted from one object to another. In the case of interactions between physical objects, a cause-effect relationship is inferred from the transmission of some physical quantity (e.g., velocity, energy, force) from the agent object (or cause object) to the patient object (or effect object).

Research in the field of intuitive physics has shown that people understand interactions between physical objects – including collisions – in terms of *transmission of impetus* (Clement, 1982; diSessa, 1993; Halloun & Hestenes, 1985; McCloskey, 1983). For instance, people intuitively understand a Michottean collision such as that depicted in Fig. 1 as an agent (i.e., object *A*) transmitting impetus to a patient (i.e., object *B*), which *resists* the transmission of impetus to a certain degree (Hubbard, 2013c; Hubbard & Ruppel, 2002; White, 2009a). As a consequence, *A* is perceived to exert a force on *B*, whereas *B* is perceived to exert a small or null force on *A* (White, 2007, 2009b). Incidentally, this is at odds with the symmetry of forces implied by Newton's Third Law, which states that the force that *A* exerts on *B* is equal and opposite to the force that *B* exerts on *A*. Hubbard (2013c, p. 642) speculated that “perhaps observers experience impressions of causality when viewing launching effect displays not because they directly perceive causality, but because behavior of the mover [object *A*] and target [object *B*] in launching effect displays match an impetus heuristic used when predicting outcomes of collision events” (see also Hubbard & Ruppel, 2002). This hypothesis appears to dismiss the possibility of direct visual impressions of causality, but to do so would be at odds with recent findings supporting the existence of such low-level visual impressions (e.g., Buehner & Humphreys, 2010; Kim et al., 2013; Moors et al., 2017; Rolfs et al., 2013; Scholl & Nakayama, 2004). A hypothesis compatible with these recent findings is that visual impressions of launching are impervious to the impetus transmission heuristic because they result from a visual module (Leslie & Keeble, 1987; Scholl & Tremoulet, 2000; see also Firestone & Scholl, 2016); however, *explicit* reports of causality would be primarily driven by high-level interpretations of the stimuli based on the impetus transmission heuristic, rather than by visual impressions of launching. The hypothesis that we set out to test in this study is that explicit reports of causality are primarily driven by people's intuitive understanding of the physics of collision, rather than by genuine visual impressions. This is consistent with the idea that explicit causal responses are based more on an intuitive understanding of the physical situation (i.e., on high-level cognitive processes) than on low-level perceptual cues (Schlottmann, 2000, 2001).

We tested the hypothesis in two distinct stimulus conditions: in Experiment 1, we presented participants with simulated collisions involving depictions of realistic spheres differing in implied masses, whereas in Experiment 2, we presented participants with simulated collisions involving depictions of non-material spheres. We speculated that explicit reports of causality could be primarily driven by people's intuitive understanding of the physics of collision when the stimuli are sufficiently similar to real life physical collisions (Experiment 1), because this is the domain to which intuitive physics of collisions normally applies. We also speculated that, when the stimuli lack the material properties of real life objects as in Experiment 2, participants' responses could be driven by a different source of information, namely visual impressions of launching.

The intuitive understanding of collisions and the explicit reports of causality are two distinct constructs that should be measured

---

[1] Within the frame of the “representational momentum” paradigm, Hubbard suggested that the remembered vanishing location of *B* also constitutes an indirect measure of perceptual causality (see Hubbard, 2013c; Hubbard & Ruppel, 2002). This claim, however, has been called into question by Choi and Scholl (2006).

separately. Consistently with previous studies (Kaiser & Proffitt, 1987; Sanborn, Mansinghka, & Griffiths, 2013; Schlottmann & Anderson, 1993; Vicovaro & Burigana, 2014, 2016), we measured participants' intuitive understanding of collisions by means of naturalness ratings, that is by asking them to rate the apparent naturalness of simulated collisions. In the light of the results of previous studies in the field of intuitive physics, naturalness ratings are expected to reflect the impetus transmission heuristic (but see Section 1.2 below). We measured explicit causal reports by means of ratings of causality, that is, by asking participants to indicate to what extent the motion of *B* appeared to be caused by the collision with *A* (see also Hubbard & Ruppel, 2013; Sanborn et al., 2013; Schlottmann & Anderson, 1993). If explicit causal reports are also driven by the impetus transmission heuristic – as suggested above – then naturalness and causality ratings should be strongly correlated, that is, they should vary with the same variables in a similar way. The more consistent with impetus transmission a collision appears to be, the more 'natural' it should appear to be and the more the motion of *B* should be rated as caused by the collision with *A*. A competing hypothesis is that explicit causal reports are primarily driven by genuine visual impressions of launching. As emphasized by Michotte (1963), the launching effect is independent of the observer's *knowledge* about collisions (see also Scholl & Tremoulet, 2000), and thus it should be independent of the impetus transmission heuristic. If explicit causal reports are driven by the launching effect, then there is no reason to expect that ratings of naturalness and causality will be positively associated, since it is assumed that only the former are driven by the impetus transmission heuristic.

Chapter 5 of Michotte's book (1963) argued that the launching effect is not influenced by properties of the colliding objects, such as their shape, color and dimensions. In Experiment 27, Michotte showed that the launching effect occurs even when *A* and *B* are shadows projected onto a screen, whilst in Experiment 28 he showed that the effect occurs when *A* is a real wooden sphere and *B* is a bright circle projected onto a screen. Michotte (1963) concluded that one of the keys to the launching effect is the presence of two phenomenally distinct objects, irrespective of their features. This provides strong support for the hypothesis that the launching effect is a genuine visual phenomenon, because it shows that the effect depends only on the low-level kinematic properties of the stimuli and is unaffected by stimulus properties that are presumed to be processed at a higher cognitive level – such as the implied mass of the colliding objects.[2] In contrast, dynamic properties of the colliding objects play a critical role in the impetus transmission heuristic, as people believe that impetus and resistance increase with mass (Halloun & Hestenes, 1985). For a fixed pre-collision velocity of *A*, the larger the mass of *A*, the greater the impetus which is believed to be transmitted from *A* to *B*, and the larger the mass of *B*, the greater *B*'s resistance to impetus transmission is believed to be. The crucial point is that if explicit causal reports of Michottean collisions are based on visual perception of the launching effect, then causality ratings should be independent of the implied masses of *A* and *B*. If, however, explicit causal reports are based on the impetus transmission heuristic, then the implied masses of *A* and *B* should exert approximately the same influence on ratings of causality and on ratings of naturalness. In Experiment 1 we tested these predictions by directly comparing naturalness and causality ratings for Michottean collisions between simulated colliding objects differing in implied mass.

1.2. *A revision of the impetus transmission heuristic*

According to the impetus transmission heuristic, an agent transmits impetus to a patient, the patient resists this to some extent and, most importantly, the patient does not transmit any impetus to the agent (diSessa, 1993; McCloskey, 1983; White, 2009a). In Michottean collisions, the role of agent is assigned to the object that moves first (i.e., *A*), whereas the role of patient is assigned to the initially stationary object (i.e., *B*). The force that *A* is perceived to exert on *B* is larger than the force that *B* is perceived to exert on *A*, which is inconsistent with Newton's Third Law, but consistent with impetus transmission heuristic (White, 2007, 2009b). However, some deviations from the predictions of the impetus transmission heuristic have also been reported. Hubbard and Ruppel (2013) presented participants with various kinematic patterns of colliding objects, including a case in which object *A* appeared to shatter on impact with *B*, which remained stationary after the collision. This gives the impression of a relatively light or fragile object that is smashed when it collides with a very massive object, for example a glass bottle shattering when it hits a rock. This shattering-type stimulus produced higher ratings of force and of causality for the patient (i.e., *B*) than for the agent (i.e., *A*), showing that – inconsistently with the original impetus transmission heuristic – people may sometimes believe that the patient can exert a large amount of force on the agent. Similarly in some of our everyday life experiences - for instance, when we punch a wall, jump on a rubber mattress or drop a superball on the floor - the patient does not simply 'resist' impetus transmission, it also appears to impart a force to the agent. In other words, force can be transmitted from patient to agent as well as from agent to patient. The force that the agent appears to transmit to the patient can be attributed to impetus, whereas the force that the patient appears to transmit to the agent can be attributed to a 'reaction' of the patient to impetus transmission. We call this revised impetus heuristic the 'bidirectional force transmission heuristic' (BFTH). As compared with the impetus transmission heuristic, the BFTH does share the general assumption that people's intuitive understanding of physics mostly relies on heuristic processes, rather than on internalized physical laws (cf. Sanborn et al., 2013; see Section 4.2 for a discussion). However, the BFTH suggests that people's intuitive understanding of collisions is more consistent with Newton's Third Law than what it has been suggested in previous accounts of the impetus transmission heuristic (Clement, 1982; diSessa, 1993; Halloun & Hestenes, 1985; Hubbard, 2013c; Hubbard & Ruppel, 2002; McCloskey, 1983; White, 2009a).

According to the BFTH, Michottean collisions may involve 'backwards' transmission of force from *B* to *A* as well as 'onwards' transmission of force from *A* to *B*. The amount of backwards force transmission should increase with the resistance of *B*, in other words, the larger *B*'s resistance, the greater its tendency to transmit a force to *A*. In turn, *A* may resist this force. The implication is that if *B* exerts very high resistance it will remain stationary despite the force transmitted by *A*, whereas if *A* exerts very high resistance it will maintain its pre-collision motion despite the force transmitted by *B*. In Experiment 1 of this study, we held the pre-collision velocity of *A* constant and manipulated the implied masses of *A* and *B* through manipulations of the material from which they appeared to be made. The BFTH allows one to make a number of predictions about this experimental situation and these are schematized in Fig. 2. For a fixed pre-collision velocity of *A*, the post-collision velocity of *B* should increase with the mass of *A* (i.e., with the force transmitted by *A*) and it should decrease with the mass of *B* (i.e., with the resistance of *B* to the force transmitted by *A*). The post-collision velocity of *A* should also increase with the mass of *A* (i.e., with the resistance of *A* to the force transmitted by *B*) and it should decrease with the mass of *B* (i.e., with the amount of force transmitted by *B*). For instance, when the implied mass of *A* greatly exceeds the implied mass of *B*, collisions should be rated as more natural if *A* keeps on moving in the same direction after the collision and if the post-collision velocity of *B* is relatively high. Conversely, when the implied mass of *B* greatly

---

[2] Some authors have put this strong claim into question, suggesting that non-kinematic properties of the scene that are processed at a cognitive level may affect the visual impression of launching. For instance, Young and Falmier (2008) argued that a predictive color change of *A* can be perceived as the cause of the motion of *B*. However, Choi and Scholl (2006) argued against this type of top-down effect in the visual perception of causality, suggesting that the influence of properties of the scene that are processed at a high cognitive level on explicit causal responses probably reflects causal *thinking*, rather than causal *perception* (see also Firestone & Scholl, 2016).

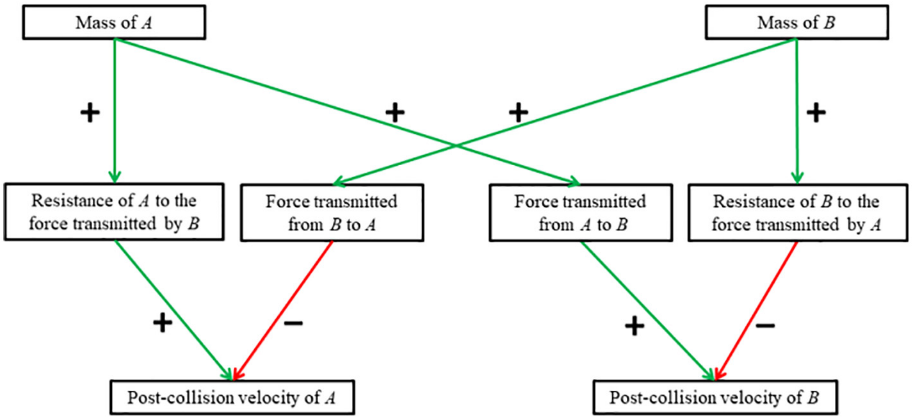

Fig. 2. A schema of the predictions of the bidirectional force transmission heuristic (BFTH) for Experiment 1. Green arrows and symbol '+' stand for 'increases with', red arrows and symbol '−' stand for 'decreases with.' (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

exceeds the implied mass of *A*, collisions should be rated as more natural if *A* bounces back from the collision and if the post-collision velocity of *B* is relatively low. The present account of the BFTH allows us to make general, *qualitative* predictions about participants' understanding of the effects of implied masses on post-collision velocities, whereas it does not afford fine-grained *quantitative* predictions. For instance, participants are expected to intuitively understand that the post-collision velocities of *A* and *B* increase with increases in the mass of *A* and with decreases in the mass of *B*; however, participants are not expected to have accurate quantitative knowledge of the range of post-collision velocities that would be 'natural' for a given combination of implied masses of *A* and *B*. In this regard, it is perhaps worth noting that, like the BFTH, most models and theories that rely on the idea that people's understanding of physics is based on heuristic processes make qualitative predictions about participants' responses (e.g., diSessa, 1993; McCloskey, 1983; Schultz et al., 1986; White, 2009a; cf. Sanborn et al., 2013).

2. Experiment 1

2.1. Participants

Forty psychology students at the University of Padua participated in the experiment on a voluntary basis. They all had normal or corrected-to-normal vision and were aged from 19 to 30 years ($M = 22.95$ years, $SD = 2.77$ years), 28 were women and 12 were men. On average, they had studied physics at school for 2.88 years ($SD = 1.60$ years). All were naive to the purposes of the experiment. Prior to the experiment, participants read and signed a consent form approved by the local ethics committee (Department of General Psychology, University of Padua).

2.2. Stimuli and design

The stimuli were presented on a personal computer equipped with a 37.5 × 30 cm CRT screen and a keyboard. The participants sat about 50 cm from the screen and stimuli were presented against a black background. At the beginning of each animation a simulated sphere (*B*) was presented in the center of the screen; 60 ms later another sphere (*A*) appeared from the left edge of the screen. The apparent size of both spheres, computed from the image diameter (2.52 cm), was 8.4 cm^3 and they subtended a visual angle of about 2.88°. The two spheres were presented in the middle of the screen's y-axis, with their centers aligned horizontally. Throughout the animation one or both spheres moved uniformly along the screen's x-axis, without rotation. During the pre-collision phase, *A* moved from left to right - towards *B* - with a velocity (hereafter u_A) of 21.4 cm/s (24.26°/s), whilst *B* remained stationary.[3] During the post-collision phase the post-collision velocity of *B* (hereafter v_B) could take one of four possible values (i.e., $0.5u_A$, u_A, $2u_A$, or $4u_A$), whereas the post-collision velocity of *A* (hereafter v_A) could take one of five possible values (i.e., $-2u_A$, $-u_A$, $-0.5u_A$, 0, $0.5u_A$). Negative velocity is motion from right to left, as when *A* bounces back from the collision. Each animation lasted 1700 ms (the post-collision phase lasted 800 ms). Fig. 3 depicts one of the collision stimuli.

We manipulated the material (polystyrene; wood) from which the spheres appeared to be made. The spheres were created with 3D Studio Max. Photographic textures depicting the materials from which they appeared to be made were attached to the spheres' surfaces, and their reflectance was regulated in order to increase the resemblance to a real sphere made from the relevant material (see the below and Fig. 3). The resulting spheres appeared to be made from a light brown wood or from an off-white polystyrene. In condition $m_A < m_B$ *A* appeared to be made from polystyrene and *B* from wood, and vice versa in condition $m_A > m_B$. In condition $m_A = m_B$ both spheres appeared to be made of wood. Before and during the experiment participants were allowed to touch and grasp real spheres made of polystyrene and wood in order to facilitate identification of the materials from which the virtual spheres appeared to be made. The masses of the real spheres were 5 g and 55 g and their diameter was 4.5 cm. The participants viewed 180 experimental stimuli - the product of a 3 (relative implied mass) × 5 (v_A) × 4 (v_B) × 3 (replication) factorial design.

2.3. Procedure

Participants were randomly divided in two groups of 20 participants each. One group received the 'naturalness rating' instructions, whereas the other group received the 'causality rating' instructions. The resulting groups shared the same gender distribution (six men in both groups) and had similar mean ages (23.15 years and 22.75 years, respectively). Both groups were informed that they would see a simulation of two colliding spheres, made of polystyrene or wood, and then given an opportunity to handle the corresponding real spheres. After that, they were given the appropriate set of written instructions. Both groups viewed the same set of 180 experimental stimuli as described above.

The instructions for the naturalness rating group were as follows: "Your task is to rate the naturalness of each collision using a number from 0 (completely unnatural) to 100 (completely natural). Completely

---

[3] In order to keep the number of stimuli within reasonable limits, in Experiments 1 and 2 we presented the participants with stimulus collisions in which the initially moving object always moved from left to right (see Fig. 3). Although this methodological choice is common in causal perception studies, it prevented us from testing the possible effects of the direction of the pre-collision motion on participants' responses. These effects did not emerge in the few causal perception studies in which the direction variable was systematically manipulated (e.g., Scholl & Nakayama, 2002; White & Milne, 1997).

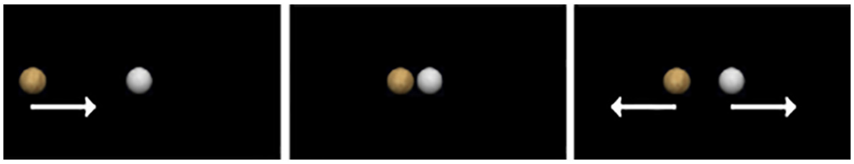

Fig. 3. Three frames of a stimulus collision between a simulated wooden sphere (left) and a simulated polystyrene sphere (right). Arrows were added to indicate which objects are moving in the three stages of the collision event.

natural (100) means that after the collision both spheres appear to move in a way that is consistent with the laws of physics and thus the collision looks very similar to a real collision. Completely unnatural (0) means that after the collision the two spheres appear to move in a way that violates the laws of physics and thus the collision looks very different from a real collision. When you are judging the naturalness of each collision you should use the number 0 only if the collision looks totally inconsistent with physical laws, and the number 100 only when the collision looks totally consistent with physical laws. If a collision looks neither totally consistent nor totally inconsistent with physical laws you should use a number between 1 and 99. The more natural a collision looks, the larger the number you should use. You can use any whole number up to and including 100, for example 20, 52, or 95."

The instructions for the causal rating group were as follows: "Your task is to rate the extent to which the movement of the central sphere (initially stationary) appears to have been caused by the collision with the left sphere (initially moving). After watching the collision you should give it a number between 0 and 100, where 0 corresponds to "the motion of the central sphere was not caused at all by the collision with the left sphere", and 100 corresponds to "the motion of the central sphere was completely caused by the collision with the left sphere". When judging the degree to which the left sphere is responsible for the motion of the central sphere, you should use the number 0 only when the motion of the central sphere appears to be totally independent of the collision (not caused by it at all), in other words when the central sphere appears to have started moving because of some internal force rather than because of the collision with the left sphere. You should use the number 100 only when the motion of the central sphere appears to be solely due to the collision. If the motion of the central sphere looks neither totally independent of, nor totally due to the collision with the left sphere you should use a number between 1 and 99. The more the central sphere's motion appears to have been caused by the collision with the left sphere, the larger the number you should use. You can use any whole number up to and including 100, for examples 20, 52, or 95. Remember that both spheres can move after the collision, and that your task is to judge the extent to which the motion of the *central* sphere appears to have been caused by the collision with the left sphere."

Participants in both groups were then informed that they could watch each collision as many times as they wanted by pressing the spacebar on the keyboard. When they felt ready to respond, then they should press the return key, type the appropriate number and then press Return again. After reading the instructions, the participants watched ten randomly chosen stimuli to familiarize them with the task. It is worth emphasizing that the participants in the naturalness rating group were asked to focus on the post-collision motions of both spheres, which is consistent with studies on the intuitive physics of collisions in which the post-collision velocities of both *A* and *B* were systematically manipulated (Kaiser & Proffitt, 1987; Vicovaro & Burigana, 2016). However, the participants in the causality rating group were asked to focus on the post-collision motion of *B*, which is consistent with studies on causal judgments in which, differently from Experiment 1, *A* is typically kept stationary during the post-collision phase (e.g., Sanborn et al., 2013; Schlottmann & Anderson, 1993).

2.4. Results and discussion

Fig. 4 shows the mean naturalness ratings (top panels) and the mean causality ratings (bottom panels) for each group. Visual inspection of the figure suggests that the effects of v_B (horizontal axis) and v_A (separate lines) on both naturalness and causality ratings were strongly mediated by the relative implied mass of the spheres (separate graphs).

2.4.1. Naturalness ratings

We performed a three-way within-participants ANOVA in naturalness ratings with the factors v_A, v_B, and relative implied mass. There were main effects of all three factors (v_A: $F(4,76) = 25.16$, $p < .001$, $\eta_p^2 = 0.57$; v_B: $F(3,57) = 24.2$, $p < .001$, $\eta_p^2 = 0.56$; relative implied mass: $F(2,38) = 40.15$, $p < .001$, $\eta_p^2 = 0.68$).The two-way interactions ($v_A \times v_B$; $v_A \times$ relative implied mass; $v_B \times$ relative implied mass) were also statistically significant ($F(12,228) = 6.77$, $p < .001$, $\eta_p^2 = 0.26$; $F(8,152) = 20.82$, $p < .001$, $\eta_p^2 = 0.52$; $F(6,114) = 19.32$, $p < .001$, $\eta_p^2 = 0.50$ respectively). The three-way interaction was also statistically significant, $F(24,456) = 3.712$, $p < .001$, $\eta_p^2 = 0.20$. We explored the effects of v_B and v_A on the naturalness ratings, and their consistency with the BFTH predictions using a series of post hoc Tukey's HSD tests (with $\alpha = 0.05$) on each level of relative implied mass. For the sake of simplicity, we explored the effects of v_B and v_A separately, by comparing the marginal means of the levels of each factor with each other.

Condition $m_A < m_B$. According to the BFTH, when the mass of *B* is large and the mass of *A* is small – as in the $m_A < m_B$ condition (top left panel in Fig. 4) – collisions where the values of v_B and v_A are relatively small should be rated most natural, because the BFTH predicts that the post-collision velocities of both *B* and *A* should decrease with the mass of *B* and increase with the mass of *A* (see Fig. 2). Consistent with this prediction, post hoc tests showed that the naturalness ratings (averaged over v_A) decreased with v_B (i.e., $v_B = 0.5u_A$ was rated significantly more natural than $v_B = u_A$, which in turn was rated significantly more natural than $v_B = 2u_A$, etc.). In contrast, no statistically significant difference emerged between the naturalness ratings for the five levels of v_A (averaged over v_B), but also consistent with the BFTH predictions, the mean naturalness ratings were higher for collisions in which $v_A < 0$, that is, collisions in which *A* bounced back from the collision.

Condition $m_A > m_B$. According to the BFTH, when the mass of *A* is large and the mass of *B* is small – as in the $m_A > m_B$ condition (top right panel in Fig. 4) – collisions in which the values of v_B and v_A are relatively large should be rated most natural, because the BFTH predicts that the post-collision velocities of *B* and *A* should increase with the mass of *A* and decrease with the mass of *B* (see Fig. 2). Consistent with this prediction, post hoc tests showed that the naturalness ratings (averaged over v_A) for $v_B = u_A$ and $v_B = 2u_A$ were significantly higher than those for $v_B = 0.5u_A$. The naturalness ratings for $v_B = 2u_A$ were also significantly higher than those for $v_B = 4u_A$, and the naturalness ratings for the pairs $v_B = u_A$ and $v_B = 2u_A$, $v_B = u_A$ and $v_B = 4u_A$, $v_B = 0.5u_A$ and $v_B = 4u_A$ were not significantly different one another. As regards the effects of v_A, post hoc tests showed that, consistent with the BFTH predictions, the naturalness ratings (averaged over v_B) for $v_A = 0.5u_A$ (the highest level of v_A) were significantly higher than for the other levels of v_A, whereas the naturalness ratings for $v_A = -2u_A$ (the lowest level of v_A) were significantly lower than for the other levels of v_A. None of the other differences between the levels of v_A reached statistical significance.

Condition $m_A = m_B$. In this condition (top central panel in Fig. 4) we would expect a pattern of results somewhere in between the results for conditions $m_A < m_B$ and $m_A > m_B$. Post hoc tests showed that the naturalness ratings (averaged over v_A) for $v_B = 4u_A$ were significantly

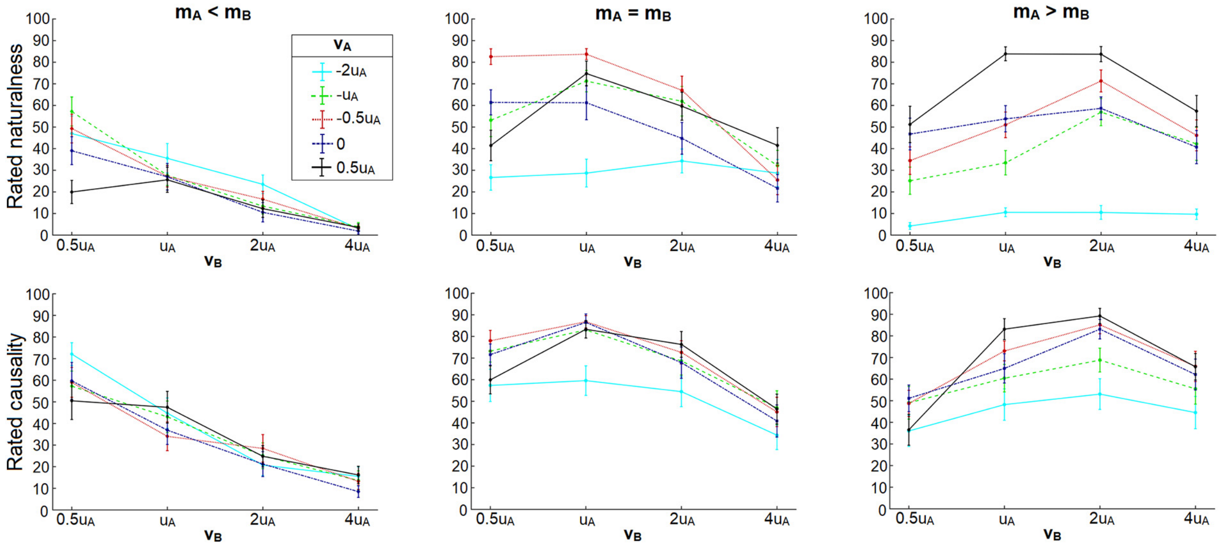

Fig. 4. The mean rated naturalness (top panels) and the mean rated causality (bottom panels) from Experiment 1 represented as a function of v_B (horizontal axis), v_A (separate lines), and relative implied mass (separate panels). The leftmost panels are for condition $m_A < m_B$, the central panels are for condition $m_A = m_B$, and the rightmost panels are for condition $m_A > m_B$. In all panels, the vertical bars represent the standard errors of the means.

lower than for the other levels of v_B, which did not differ from each other. Moreover, the naturalness ratings (averaged over v_B) for $v_A = -2u_A$ were significantly lower than for the other levels of v_A, and those for $v_A = -0.5u_A$ were significantly higher than those for $v_A = 0$. None of the other differences between the levels of v_A reached statistical significance.

Overall, the results provide support to the hypothesis that participants intuitively understood that the post-collision velocities of *A* and *B* increased with the mass of *A* and decreased with the mass of *B*, which is consistent with the BFTH predictions. It is however worth noting that, in each of the three 'relative implied mass' conditions, many of the post-hoc comparisons between the five levels of v_A, and many of the post-hoc comparisons between the four levels of v_B, did not reach a statistically significant level (e.g., in the $m_A < m_B$ condition, none of the post-hoc comparisons between the five levels of v_A reached a statistically significant level). In other words, in each of the three 'relative implied mass' conditions, several levels of v_A and v_B received similar ratings of naturalness. This suggests that participants lacked clear-cut quantitative knowledge of the relationship between the specific combinations of implied masses of the simulated spheres and their post-collision velocities (see also Section 4.2).

2.4.2. Causality ratings

Comparing the top and bottom panels of Fig. 4 reveals a remarkable similarity between the patterns of the naturalness and causality ratings. Analysis of the correlations between the mean naturalness ratings and the mean causality ratings for the 60 combinations of levels of the three experimental factors yielded $r = 0.89$, indicating that collisions that were rated natural were also considered to represent a causal relationship between the motions of *A* and *B*. This finding provides support for the hypothesis that causality ratings were driven by participants' intuitive understanding of the physics of collisions, that is, by the BFTH, and not by visual perception of the launching effect. We performed a three-way ($v_A \times v_B \times$ relative implied mass) within-participants ANOVA on the causality ratings. There were main effects of all three factors (v_A: $F(4,76) = 6.88$, $p < .001$, $\eta_p{}^2 = 0.27$; v_B: $F(3,57) = 19.2$, $p < .001$, $\eta_p{}^2 = 0.50$; relative implied mass: $F(2,38) = 37.97$, $p < .001$, $\eta_p{}^2 = 0.67$) and the two-way interactions ($v_A \times v_B$, $v_A \times$ relative implied mass, $v_B \times$ relative implied mass) were also statistically significant ($F(12,228) = 3.97, p < .001, \eta_p{}^2 = 0.17$; $F(8,152) = 6.08$, $p < .001$, $\eta_p{}^2 = 0.24$; $F(6,114) = 27.01$, $p < .001$, $\eta_p{}^2 = 0.59$, respectively). The three-way interaction was not statistically significant, $F(24,456) = 0.79, p = .746, \eta_p{}^2 = 0.04$. As with the naturalness ratings we conducted a series of post hoc Tukey's HSD tests (with $\alpha = 0.05$) for each level of relative implied mass, comparing the marginal means of for all levels of factors v_B and v_A separately.

Condition $m_A > m_B$. One of the most important launching effect variables is the ratio between the pre-collision velocity of *A* and the post-collision velocity of *B*. Michotte (1963) showed that a clear impression of launching is usually obtained if the post-collision velocity of *B* is less than twice the pre-collision velocity of *A*, that launching is perceived about 50% of the time when the post-collision velocity of *B* is approximately twice the pre-collision velocity of *A*, and that launching is not perceived when the post-collision velocity of *B* is greater than twice the pre-collision velocity of *A*. On this basis the stimuli for Experiment 1 were *optimal* for visual perception of the launching effect (i.e., $v_B = 0.5u_A$ and $v_B = u_A$), *suboptimal* (i.e., $v_B = 2u_A$) or *not conducive* to it (i.e., $v_B = 4u_A$). The bottom right panel of Fig. 4 shows a striking discrepancy between the results of Experiment 1 and Michotte's findings. Post hoc tests on the causality ratings for the $m_A > m_B$ condition (averaged over v_A) showed that collisions that were suboptimal or not conducive to visual perception of the launching effect (i.e., $v_B = 2u_A$ and $v_B = 4u_A$) received significantly higher causality ratings than one of the two groups of collisions that were optimal for perception of the launching effect (i.e., $v_B = 0.5u_A$). This is probably because, in accordance with the BFTH, the participants expected the post-collision velocity of *B* to be relatively high when $m_A > m_B$ (see Fig. 2 and the top right panel of Fig. 4). Post hoc tests also showed that the causality ratings for $v_B = u_A$ were significantly higher than for $v_B = 0.5u_A$, and that the causality ratings for $v_B = 2u_A$ were significantly higher than for $v_B = 4u_A$. None of the other differences between the levels of v_B reached statistical significance. Post hoc tests also showed that the causality ratings (averaged over v_B) for $v_A = -2u_A$ were significantly lower than for the other levels of v_A (except $v_A = -u_A$), and that none of the other differences between the levels of v_A reached statistical significance.

Condition $m_A < m_B$. In this condition, the pattern of causality ratings (bottom left panel of Fig. 4) was practically identical to the pattern of the corresponding naturalness ratings (top left panel). Indeed, post hoc tests showed that causality ratings (averaged over v_A) decreased with v_B (i.e., $v_B = 0.5u_A$ received significantly higher causality ratings than $v_B = u_A$, which in turn received significantly higher causality ratings than $v_B = 2u_A$, etc.), and post hoc tests did not reveal any significant differences between the causality ratings for the five levels of v_A

(averaged over v_B), although in descriptive terms the mean causality ratings were higher for collisions in which *A* bounced back from the collision (i.e., $v_A < 0$).

Condition $m_A = m_B$. The pattern of causality ratings (bottom central panel of Fig. 4) was similar to the pattern of the corresponding naturalness ratings (top central panel). Indeed, the causality ratings (averaged over v_A) for $v_B = 4u_A$ were significantly lower than for the other levels of v_B, and the causality ratings (averaged over v_B) for $v_A = -2u_A$ were significantly lower than for the other levels of v_A. Post hoc tests also revealed that the causality ratings for $v_B = u_A$ were significantly higher than for $v_B = 0.5u_A$, which is at odds with Michotte's findings because both types of collision should be optimal for the launching effect. None of the other differences between the levels of v_A or v_B reached statistical significance. Therefore, collisions that were suboptimal for the launching effect (i.e., $v_B = 2u_A$) did not receive significantly lower causality ratings than collisions that were optimal for the launching effect (i.e., $v_B = 0.5u_A$ and $v_B = u_A$), which is also at odds with Michotte's findings.

The effects of v_A on causality ratings are worthy of comment. As Michotte (1963) argued, the post-collision velocity of *A* has small influence on the launching effect, provided that it does not exceed the post-collision velocity of *B* (otherwise *A* would be perceived to overtake *B*). In particular, Michotte showed that, when the other stimulus conditions were conducive to visual perception of the launching effect, the phenomenon occurred independently of whether, during the post-collision phase, *A* kept on moving in the same direction as before the collision or bounced back.[4] If causality ratings were driven by the launching effect, then v_A should have had a negligible effect on them. The vertical separation of the curves in the bottom panels in Fig. 4 and the results of the statistical analysis clearly show that this was not the case. Post hoc tests showed that $v_A = -2u_A$ received significantly lower causality ratings than the other levels of v_A when $m_A > m_B$ and when $m_A = m_B$. Conversely, $v_A = -2u_A$ received the highest mean causality rating when $m_A < m_B$. To our knowledge, this is the first evidence that the effects of the post-collision velocity of *A* on causality ratings are strongly mediated by the relationship between the implied masses of the colliding objects.

One may note that v_A generally had a smaller effect on causality ratings than on naturalness ratings; this is reflected in the fact that in Fig. 4 the vertical separation of the curves is smaller in the bottom panels than the top panels. This finding is hardly surprising, however, because whilst the naturalness rating group was instructed to base its ratings on the post-collision motion of *both* spheres, the causal rating group was explicitly instructed to base its ratings on the extent to which *the motion of B* appeared to be caused by the collision with *A*. In fact, it is quite surprising that despite these instructions v_A had an effect on causality ratings (albeit a relatively small one). This indicates that the causal relationship between the pre-collision motion of *A* and the post-collision motion of *B* was judged to be weaker when the post-collision motion of *A* was inconsistent with the BFTH.

3. Experiment 2

The results of Experiment 1 showed that explicit reports of causality were driven by the BFTH, rather than by visual impressions of launching. This does not exclude the possibility that some of the stimuli used in Experiment 1 elicited genuine visual impressions of launching, but it suggests that explicit reports were overwhelmingly influenced by the intuitive physics of collisions. One possibility is that the use of the BFTH was prompted by the perceived realism of the colliding objects, by the fact that they had definite implied masses, that these implied masses were systematically manipulated, and that participants could haptically perceive the masses of corresponding real spheres. Arguably, the BFTH provides people with representations of real life physical events that obviously involve real material objects. It can be hypothesized, therefore, that the BFTH plays a role in explicit reports of causality only when the simulated events are sufficiently similar to real life physical events, because this is the domain to which the BFTH presumably applies. In the vast majority of the experiments on the perception of causality, observers have been presented with simulated collisions between abstract ‘non-material’ shapes which provide no perceptual or cognitive cues to their implied masses (see Fig. 1). If the use of the BFTH to arrive at ratings of causality in Experiment 1 was due to the features of the stimuli and/or of the experimental procedure and design, then one would predict that the BFTH would not influence causal reports in relation to collisions between unrealistic, non-material objects, in which the implied masses of the colliding objects are not manipulated. In such conditions, individuals might be expected to base their causality ratings on another source of information, namely visual impressions of launching.

In Experiment 2, we tested this hypothesis by comparing naturalness and causality ratings for simulated collisions between non-material colliding spheres. Manipulations of the velocities of the colliding objects were the same as in Experiment 1, but all the spheres had a uniform greenish surface and there were no cues to their material or mass, although their shading was regulated in order to give them a clear 3-D appearance. If ratings of causality are driven by the launching effect, then v_A should have a small or null effect, because the launching effect is largely independent of it (see Section 2.4.2). Additionally, participants should use the upper end of the rating scale for collisions that are conducive to the launching effect (i.e., $v_B = 0.5u_A$ and $v_B = u_A$), the middle part of the scale for collisions that are suboptimal for the launching effect (i.e., $v_B = 2u_A$), and the lower end of the scale for collisions that are not conducive to the launching effect (i.e., $v_B = 4u_A$). It is more difficult to predict naturalness ratings as participants are explicitly asked to evaluate the physical plausibility of the collision and there does not appear to be a viable alternative to the BFTH. One possibility is that participants use the BFTH, making the assumption that *A* and *B* have the same implied mass because they appear identical. If this is the case, then the naturalness ratings should be similar to those obtained in the $m_A = m_B$ condition of Experiment 1 (top central panel in Fig. 4).

3.1. Participants

Forty psychology students at the University of Padua participated in the experiment on a voluntary basis. They all had normal or corrected-to-normal vision and were aged between 19 and 33 years ($M = 22.85$ years, $SD = 2.84$ years), 26 were women and 14 were men. On average they had studied physics at school for 2.95 years ($SD = 1.36$ years). All were naive to the purposes of the experiment. None of them had participated in Experiment 1. Prior to the experiment participants read and signed a consent form approved by the local ethics committee (Department of General Psychology, University of Padua).

3.2. Stimuli and design

The stimuli and the design were the same as in Experiment 1, except that the simulated spheres (created with 3D Studio Max) all had a smooth, 3-D greenish appearance. The participants were presented with 60 experimental stimuli, resulting from a 5 (v_A) × 4 (v_B) × 3 (replication) factorial design.

---

[4] An exception to this general rule is when $v_A = -v_B$, that is, when *A* bounces back with a velocity which is the opposite of the post-collision velocity of *B*. In this case the visual system would tend to group the symmetrical post-collision motions of *A* and *B*, thus disrupting the launching effect (Michotte, 1963, Experiment 25). When, during the post-collision phase, *A* kept on moving exactly with the same velocity of *B*, the launching effect left place to the “entraining” effect (Michotte, 1963, Experiment 2). Like the launching effect, also the entraining effect gives rise to the impression that *A* causes the motion of *B*.

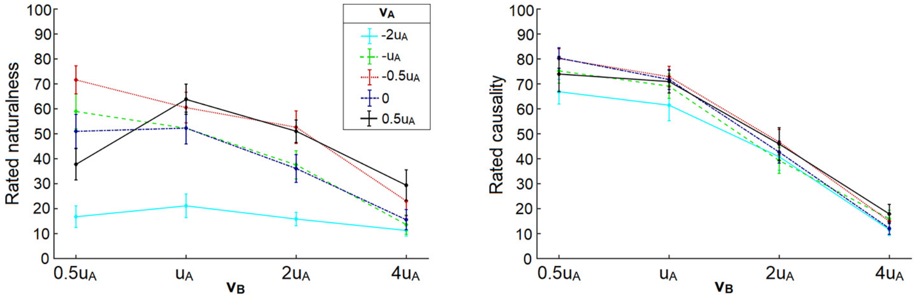

Fig. 5. The mean rated naturalness (left) and the mean rated causality (right) from Experiment 2, as a function of v_B (horizontal axis) and v_A (separate lines). The vertical bars represent the standard errors of the means.

3.3. Procedure

As in Experiment 1, the participants were randomly divided in two groups of 20 participants each. One group received the naturalness rating instructions and the other the causality rating instructions. The two groups had the same gender distribution (7 men in both groups) and similar mean ages (22.95 vs. 23.2 years, respectively). Both groups were presented with the same set of 60 experimental stimuli. The procedure and the instructions for each group were the same as in Experiment 1, except that all references to the materials of the spheres was removed and participants did not handle any real spheres before or during the experiment.

3.4. Results and discussion

Fig. 5 shows the mean naturalness ratings (left panel) and the mean causality ratings (right panel) for each group. Visual inspection of the figure reveals discrepancies between the two types of rating.

3.4.1. Naturalness ratings

We performed a two-way ($v_A \times v_B$) within-participants ANOVA on the naturalness ratings. There were main effects of v_A and v_B ($F(4,76) = 11.91$, $p < .001$, $\eta_p^2 = 0.39$; $F(3,57) = 55.91$, $p < .001$, $\eta_p^2 = 0.75$, respectively). The two-way interaction was also statistically significant, $F(12,228) = 5.65$, $p < .001$, $\eta_p^2 = 0.23$.

Comparison of the left panel of Fig. 5 and the top panels in Fig. 4 shows that the naturalness ratings in Experiment 2 were qualitatively similar to the naturalness ratings for the $m_A = m_B$ condition in Experiment 1 (simulated wooden spheres). In both cases the ordering of the mean naturalness ratings (averaged across v_A) was $v_B = u_A$, $v_B = 0.5u_A$, $v_B = 2u_A$, then $v_B = 4u_A$. Moreover, in both cases post hoc Tukey's HSD tests ($\alpha = 0.05$) showed that the naturalness ratings (averaged across v_B) for $v_A = -2u_A$ were significantly lower than for the other levels of v_A, and that those for $v_A = -0.5u_A$ were significantly higher than for $v_A = 0$. Of course, this pattern of results is clearly different from those for the $m_A < m_B$ and $m_A > m_B$ conditions in Experiment 1. These findings are consistent with the hypothesis that participants responded according to the BFTH, making the assumption that *A* and *B* had the same mass. A post hoc Tukey's HSD test ($\alpha = 0.05$) showed that the naturalness ratings (averaged across v_A) for $v_B = 4u_A$ were significantly lower than for the other levels of v_B, and that the naturalness ratings for $v_B = u_A$ were significantly higher than for $v_B = 2u_A$. The naturalness ratings for $v_B = 0.5u_A$ and $v_B = u_A$ were not significantly different from each other, as were not those for $v_B = 0.5u_A$ and $v_B = 2u_A$.

3.4.2. Causality ratings

There was a high positive correlation between the mean naturalness ratings and the mean causality ratings ($r = 0.73$), albeit lower than in Experiment 1. Despite this, further statistical analyses revealed some noteworthy discrepancies between the naturalness and causality ratings. We performed a two-way ($v_A \times v_B$) within-participants ANOVA on the causality ratings. There was no main effect of v_A, $F(4,76) = 2.11$, $p = .088$, $\eta_p^2 = 0.09$, but there was a main effect of v_B, $F(3,57) = 104.3, p < .001, \eta_p^2 = 0.85$. The two-way interaction was not statistically significant, $F(12,228) = 0.892$, $p = .556$, $\eta_p^2 = 0.04$. v_A had a clear effect on the naturalness ratings, but a small or null effect on the causality ratings, and this is reflected in the lack of vertical separation between the curves in the right panel of Fig. 5 and confirmed by the results of the ANOVA. The lack of an effect of v_A is consistent with the hypothesis that the causality ratings in Experiment 2 were primarily driven by visual impressions of launching, which are largely independent of v_A (see Section 2.4.2). Note that in the $m_A = m_B$ condition of Experiment 1 the causality ratings (averaged over v_B) for $v_A = -2u_A$ were significantly lower than those for the other levels of (see Section 2.4.2).

The effects of v_B were also consistent with the hypothesis that the causality ratings in Experiment 2 were primarily driven by genuine visual impressions. High mean causality ratings were obtained for collisions that were optimal for the launching effect ($v_B = 0.5u_A$: $M = 75.3$, $SD = 22.9$; $v_B = u_A$: $M = 69.2$, $SD = 21.5$); moderate ratings ($M = 43$, $SD = 23.6$) were obtained for collisions that were suboptimal for the launching effect (i.e., $v_B = 2u_A$); and very low causality ratings ($M = 14.47$, $SD = 13.9$) were obtained for collisions that were not conducive to the launching effect (i.e., $v_B = 4u_A$). A post hoc Tukey's HSD test ($\alpha = 0.05$) showed that the causality ratings (averaged across v_A) for $v_B = 0.5u_A$ and $v_B = u_A$ were similar and significantly higher than those for $v_B = 2u_A$ and $v_B = 4u_A$, and that causality ratings for $v_B = 2u_A$ were significantly higher than for $v_B = 4u_A$. Therefore, collisions in which $v_B = 0.5u_A$ received higher causality ratings than collisions in which $v_B = 2u_A$, although the former were not judged more natural than the latter (see Section 3.4.1). Note that, in the $m_A = m_B$ condition of Experiment 1, collisions in which $v_B = 0.5u_A$ (i.e., optimal for the launching effect) did *not* receive significantly higher causality ratings than collisions that were suboptimal for the launching effect (i.e., $v_B = 2u_A$), and they received significantly lower causality ratings than collisions in which $v_B = u_A$, which were also optimal for the launching effect. Moreover, differently from Experiment 2, in the $m_A = m_B$ condition of Experiment 1 collisions that were suboptimal for the launching effect still received relatively high causality ratings (e.g. $v_B = 2u_A$: $M = 68$, $SD = 29.2$), and collisions that were not conducive to the launching effect received moderate causality ratings (e.g. $v_B = 4u_A$: $M = 42.7$, $SD = 31.6$). To sum up, the causality ratings in Experiment 2 were more consistent with the predictions from Michotte's (1963) account of the launching effect than with the naturalness ratings in Experiment 2 or with the causality ratings in the $m_A = m_B$ condition of Experiment 1. These results provide some support to the hypothesis that different psychological processes underlie the causality ratings in Experiments 1 and 2, namely the BFTH in the former and visual impressions of launching in the latter. However, this conclusion should be taken with caution, because it relies upon relatively small differences between the naturalness and the causality ratings in Experiment 2, and relatively small differences between the causality ratings in Experiment 2 and the causality ratings in the $m_A = m_B$ condition of Experiment 1. In the light of these results, we cannot unambiguously exclude the contribution of the BFTH to explicit judgments of causality when simulated collisions involve non-material objects.

4. General discussion

The main findings from Experiments 1 and 2 can be summarized as follows. If in Experiment 1 the causality ratings were driven by visual impressions of launching, then collisions in which $v_B = 0.5u_A$ or $v_B = u_A$ would have received high causality ratings, collisions in which $v_B = 2u_A$ would have received moderate causality ratings, and collision in which $v_B = 4u_A$ would have received low causality ratings. Moreover, neither the implied masses of the spheres nor the post-collision velocity of A would have affected the causality ratings. None of these predictions was confirmed by the results. Instead, the similarity between the causality ratings and the naturalness ratings, and the consistency between the latter and the BFTH predictions (see Fig. 2), suggest that causality ratings were actually driven by the BFTH. In Experiment 2, collisions in which $v_B = 0.5u_A$ or $v_B = u_A$ received high causality ratings, collisions in which $v_B = 2u_A$ received moderate causality ratings, and collision in which $v_B = 4u_A$ received low causality ratings. Moreover, the post-collision velocity of A exerted a negligible influence on the causality ratings, and some discrepancies between the naturalness and the causality ratings were observed. The results provide some support to the hypothesis that in Experiment 2 the causality ratings were mainly driven by visual impressions of launching, rather than by the BFTH.

Recent studies have shown that causality can be automatically perceived through low-level visual processing of launching stimuli (e.g., Buehner & Humphreys, 2010; Kim et al., 2013; Moors et al., 2017; Rolfs et al., 2013; Scholl & Nakayama, 2004), which provides support to the hypothesis that visual impressions of launching are the result of a visual module (Leslie & Keeble, 1987; Scholl & Tremoulet, 2000; see also Firestone & Scholl, 2016). However there is still debate about the relative contributions of visual impressions and intuitive understanding of physical interactions between objects to explicit causal responses (Bullock et al., 1982; Hubbard, 2013c; Schlottmann, 1999, 2000, 2001; Schultz et al., 1986). These results suggest that this may depend on contextual information – specifically, on the features of the stimuli and of the experimental design. When simulated collisions involved depictions of realistic objects differing in implied masses, the motion of B was judged to have been caused by the collision with A if the collision appeared to be physically plausible, that is, consistent with the BFTH. In contrast, in the case of simulated collisions between non-material objects, causality ratings appear to be primarily driven by visual impressions of launching. A part from differences in the perceived realism of the stimuli, differences between the experimental designs and procedures may also help to explain the different patterns of results. In Experiment 1, but not Experiment 2, the implied relative mass of the stimuli was systematically manipulated, and participants were allowed to touch and grasp real spheres: this may have directed participants' attention to the relationship between the implied masses and the motion of the stimuli. One limit of the current experiments is that they do not allow separation of the contributions of the perceived realism of the stimuli and of the features of the experimental procedures and designs to the causality ratings. This issue is perhaps worthy of being explored in future studies, although it does not undermine the main outcome of Experiments 1 and 2: context (i.e., features of the stimuli, of the procedure, and/or of the experimental design) affects the response pattern in a causality rating task. This suggests that, when deciding whether to interpret explicit ratings of causality in terms of genuine visual impressions or cognitive processing, the features of the stimuli and the experimental design should be considered carefully. Participants' intuitive understanding of the physical situation should always be considered as a possible explanation of explicit reports of causality, especially when non-motion properties of the stimuli are systematically manipulated and when depictions of realistic stimuli are used. For instance, participants' descriptions of relatively complex animations involving objects varying in size, shape and color (e.g., White & Milne, 1997, 1999, 2003), are likely to depend, to a large extent, on cognitive interpretations of the stimuli that are driven by intuitive physics, rather than on genuine visual impressions (see also Choi & Scholl, 2006).

4.1. Functional roles of intuitive physics and visual impressions of launching in causal reports

The collisions of which we have direct experience involve material objects with varying masses. The results of Experiment 1 suggest that in such cases causal reports are probably driven more by our intuitive understanding of the physics of collisions than by visual impressions. This has an obvious advantage because, if it is true that visual impressions of causality are modular (Leslie & Keeble, 1987; Scholl & Tremoulet, 2000) then they remain unchanged (or they change little) throughout life. Instead, intuitive knowledge of physics becomes gradually more accurate during the earlier years of individual development (Piaget, 1970), and hence so do causal responses based on that intuitive understanding (Schlottmann, 1999; Schultz et al., 1986). The more we learn about collisions, the better we can discriminate between physically plausible and implausible collisions and the more accurately we can judge whether the motion of one object was caused solely by the motion of another object, or whether we should seek another cause for its motion. For instance, imagine that we see a car that, after being touched on the rear bumper by a bike, starts moving with the same velocity that the bike had before the collision. In this case, we may have the visual impressions that the motion of the car was caused by the impact of the bike, but our intuitive knowledge of the physical situation suggests that this could not be the case. We would conclude that our visual impression of causality is deceptive and that the motion of the car was not related to the contact with the bike (e.g., by chance, the car's driver started accelerating when the bike touched the car). This example shows that in everyday life causal responses based on intuitive physics are generally more reliable than causal responses based on visual impressions.

Schlottmann (2000, 2001) suggested that visual impressions of causality have a measurable effect only in laboratory situations, when they are artificially separated from causal inference. However, from an evolutionary viewpoint, visual impressions of causality are likely to have some function. There are at least two domains in which visual impressions of causality may play a fundamental role in causal understanding. First, when information about the mass of the colliding objects is not available or not readily perceived, as in Experiment 2 of this study or the case of collisions that are seen from a great distance. Second, in the earliest stages of development, when intuitive knowledge of physical laws is still poor, visual impressions of causality may help to give a meaningful cause-effect structure to the world. This hypothesis is supported by converging evidence from developmental and animal research showing that causal perception precedes causal reasoning in terms of ontogenetic and phylogenetic development (e.g., Caggiano, Fleischer, Pomper, Giese, & Thier, 2016; Leslie & Keeble, 1987; Mascalzoni, Regolin, & Vallortigara, 2010; Matsuno & Tomonaga, 2017; Newman, Choi, Wynn, & Scholl, 2008). Presuming that the hypothesis that visual impressions of causality are the result of a visual module is correct (Leslie & Keeble, 1987; Scholl & Tremoulet, 2000; cf. Rips, 2011), explicit causal reports appear to be driven by the output of that module only when the stimuli lack properties of real life objects. As it is shown by the results of Experiment 1, when real or realistic objects are involved in the stimulus situation, the output of the module is "overridden" by intuitive knowledge of physics (see also Schlottmann, 2000, 2001). This does not mean that visual impressions of launching are weak or absent in the case of collisions involving real or realistic objects, but that explicit post-perceptual causal reasoning, which is driven by intuitive knowledge of physics, prevails over genuine visual impressions.

4.2. Are reports of naturalness and causality based on the internalization of physical constraints?

According to the original unidirectional version of the impetus transmission heuristic (diSessa, 1993; Halloun & Hestenes, 1985; Hubbard & Ruppel, 2013; McCloskey, 1983; White, 2009a), people understand collisions such as that represented in Fig. 1 as the transmission of impetus from *A* to *B*. Object *B* offers some resistance to the force imparted by *A*, but does not transmit any force to *A*, which represents a clear conflict between the impetus transmission heuristic and Newton's Third Law. However, the results of this study provide support for a partial revision of this heuristic, suggesting that people have a representation of the fact that force can also be transmitted from *B* to *A*. The discrepancy between Newton's Third Law and people's intuitive understanding of collisions is thus less severe than previously hypothesized. Because the BFTH probably stems from everyday perceptual and motor interactions with physical objects one would expect it to be somewhat consistent with physical laws, as it should allow people to make reasonably accurate predictions; however, it is necessarily subject to the limits inherent in our perceptual and cognitive systems. Responses based on the BFTH are hypothesized to be only partially consistent with the Newtonian laws of collisions. Below we test this hypothesis by evaluating the consistency between participants' responses in Experiment 1 and the Newtonian laws of collisions.

The following equations specify the post-collision velocities of the spheres (v_A and v_B) as a function of the pre-collision velocity of *A* (u_A), of their masses (m_A and m_B), and what is referred to as the coefficient of restitution (*C*):

$$v_A = u_A\,(m_A - Cm_B)/(m_A + m_B), \tag{1}$$

$$v_B = u_A\,(m_A + Cm_A)/(m_A + m_B). \tag{2}$$

Newtonian mechanics implies that $0 \le C \le 1$. According to Eq. (1), v_A increases with m_A and decreases with m_B. Moreover, assuming $C = 1$, Eq. (1) implies that $v_A > 0$ when $m_A > m_B$, which implies that *A* keeps on moving in the same direction as before the collision if its mass exceeds that of *B*. The same equation implies that *A* bounces back from the collision when $m_A < m_B$. Overall, the results of Experiment 1 are qualitatively consistent with these predictions, as when $m_A > m_B$ the highest mean ratings of naturalness and causality were obtained for $v_A = 0.5u_A$ (the highest level of v_A), whereas when $m_A < m_B$ the highest mean ratings of naturalness and causality were obtained for $v_A = -2u_A$ (the lowest level of v_A). However, Eq. (1) imposes some strict quantitative constraints on the values of v_A. For instance, if m_B is infinitely larger than m_A, then $v_A = -u_A$, which means that *A* cannot bounce back with a velocity exceeding the absolute value of $-u_A$. However, Experiment 1 showed that in the $m_A < m_B$ condition $v_A = -2u_A$ received the highest mean naturalness and causality ratings; in other words collisions with impossible values of u_A were judged to be the most natural. According to Eq. (2), v_B increases with m_A and decreases with m_B, which is consistent with the fact that $v_B = 2u_A$ received the highest mean ratings of naturalness and causality when $m_A > m_B$, whereas $v_B = 0.5u_A$ received the highest mean ratings of naturalness and causality when $m_A < m_B$. However, Eq. (2) implies that v_B can never exceed $2u_A$, as $v_B = 2u_A$ when m_A is infinitely larger than m_B. The results of Experiment 1 showed that in the $m_A > m_B$ condition $v_B = 4u_A$ received moderate mean ratings of naturalness (range 40.7–57.35, excluding $v_A = -2u_A$) and causality (range 55.5–65.85, excluding $v_A = -2u_A$), which suggests that participants were somewhat insensitive to the violation of the quantitative constraints imposed by Eq. (2).

In Experiment 1, participants were allowed to touch and grasp real polystyrene and wooden spheres weighing 5 g and 55 g, respectively. Substituting these mass values in Eq. (1) we can obtain 'plausible' v_A values, that is values of v_A that solve Eq. (1) for possible values of parameter *C* (which may vary between 0 and 1). In the $m_A < m_B$ condition, $m_A = 5$ g and $m_B = 55$ g imply plausible v_A values comprised between $-0.83u_A$ and $0.08u_A$, because $v_A = 0.08u_A$ when $C = 0$, and $v_A = -0.83u_A$ when $C = 1$. The results showed that naturalness and causality ratings (averaged over v_B) for implausible v_A values (i.e., $v_A = -2u_A$ and $v_A = 0.5u_A$) were not significantly lower than those for plausible (or nearly plausible) v_A values (i.e., $v_A = -u_A$, $v_A = -0.5u_A$, $v_A = 0$). In the $m_A = m_B$ condition, $m_A = 55$ g and $m_B = 55$ g imply plausible v_A values ranging between 0 and $0.5u_A$. The results showed that naturalness and causality ratings (averaged over v_B) for the implausible $v_A = -2u_A$ were significantly lower than those for the other levels of v_A. However, the results also showed that naturalness and causality ratings (averaged over v_B) for the implausible $v_A = -u_A$ and $v_A = -0.5u_A$ were not significantly lower than those for the plausible $v_A = 0$ and $v_A = 0.5u_A$. Actually, the results showed that the implausible $v_A = -0.5u_A$ was rated significantly more natural than the plausible $v_A = 0$. In the $m_A > m_B$ condition, $m_A = 55$ g and $m_B = 5$ g imply plausible v_A values comprised between $0.83u_A$ and $0.92u_A$. Therefore, the five levels of v_A were all implausible in this condition. Moderate or high mean ratings of naturalness and causality (averaged over v_B) were nonetheless obtained for $v_A = -0.5u_A$, $v_A = 0$, and $v_A = 0.5u_A$ (see the top right and bottom right panels in Fig. 4).

By substituting the mass values in Eq. (2), we can also obtain 'plausible' v_B values. In the $m_A < m_B$ condition, plausible v_B values were comprised between $0.083u_A$ and $0.17u_A$, because $v_A = 0.083u_A$ when $C = 0$, and $v_A = 0.17u_A$ when $C = 1$. Therefore the four levels of v_B were all implausible in this condition. The results showed that $v_B = 0.5u_A$ received moderate mean ratings of naturalness and causality (see the top left and bottom left panels in Fig. 4), and that the naturalness and causality ratings (averaged over v_A) decreased with v_B. In the $m_A = m_B$ condition, plausible v_B values ranged between $0.5u_A$ and u_A. The results showed that naturalness and causality ratings (averaged over v_A) for the implausible $v_B = 4u_A$ were significantly lower than those for the other levels of v_B. The naturalness ratings (averaged over v_A) for the implausible $v_B = 2u_A$ were not significantly lower than those for the plausible $v_B = u_A$ and $v_B = 0.5u_A$. The causality ratings (averaged over v_A) for the plausible $v_B = u_A$ were significantly higher than those for the also plausible $v_B = 0.5u_A$, but not significantly higher than those for the implausible $v_B = 2u_A$. Lastly, in the $m_A > m_B$ condition, plausible v_B values ranged between $0.92u_A$ and $1.83u_A$. Naturalness and causality ratings (averaged over v_A) for the plausible $v_B = u_A$ and the nearly plausible $v_B = 2u_A$ were significantly higher than those for the implausible $v_B = 0.5u_A$. However, the naturalness and causality ratings (averaged over v_A) for the implausible $v_B = 4u_A$ were not significantly lower than those for the plausible $v_B = u_A$.[5]

To sum up, the results of Experiment 1 are consistent with the hypothesis that participants' judgments of naturalness and causality were based on a heuristic process. Although they were qualitatively consistent with the Newtonian equations of collisions, they also violated the quantitative constraints imposed by such equations in several ways.

---

[5] In order to evaluate the degree of consistency between participants' responses and physical laws of collisions, in Eqs. (1) and (2) we substituted the mass values of the spheres made of polystyrene and wooden that participants could touch and grasp at the beginning of Experiment 1. It can be argued, however, that remembered (i.e., subjective) mass values might be smaller than the corresponding physical values, because remembered magnitudes are usually smaller than both perceived and physical magnitudes. However, because v_A and v_B depend on the ratio between m_A and m_B (see Eqs. (1) and (2)), if memory had the same relative effect on the remembered mass values of both objects, then memory effects would not have any impact on computed physically plausible values of v_A and v_B. These values would not vary substantially even if the effects of memory on the mass of *A* were slightly different from the effects of memory on the mass of *B*.

Sanborn et al. (2013) suggested that intuitive physics and causal reports are based on the internalization of physical laws. They presented participants with Michottean collisions between non-material squares and manipulated several variables: the delay between the collision and the start of movement of *B*, the spatial gap between *A* and *B*, and the ratio between the pre-collision velocity of *A* and the post-collision velocity of *B* (object *A* was always stationary during the post-collision phase). As in our study, half the participants were asked to evaluate the naturalness of the collision, whereas the other half were asked to evaluate the extent to which the motion of *B* appeared to be caused by the collision with *A*. The results showed substantial similarity between naturalness and causality ratings. The participants' responses in both tasks were well fitted by a Bayesian model referred to as the 'noisy Newton model'. The core of the model was a combination of the constraints imposed by Newtonian mechanics on the post-collision behavior of colliding objects, and uncertainty about the stimulus variables. The good fit of the model to the data supported the hypothesis that responses in both tasks were driven by internalized Newtonian principles. One undisputable merit of the model is that it made specific quantitative predictions about the effects of the manipulation of the experimental factors on participants' responses.

Because Eqs. (1) and (2) constitute the core of the noisy Newton model, this model predicts a relatively high degree of consistency between both equations and participants' ratings of naturalness and causality. In other words, the model would predict an effect of the manipulation of the implied masses of *A* and *B* on judgments of naturalness and causality, although Sanborn et al. (2013) did not test this prediction. The results of the current Experiment 1 suggest that participants' responses were only qualitatively consistent with the predictions from Eqs. (1) and (2), as they violated the quantitative constraints imposed by such equations in several ways. Likewise, in a recent study on the intuitive physics of collisions, Vicovaro and Burigana (2016) showed that participants intuitively understood that collisions between relatively elastic objects (e.g., tennis balls) implied a higher coefficient of restitution (i.e., parameter *C* in Eqs. (1) and (2)) than collisions between relatively inelastic objects (e.g., terracotta spheres). However, the results also showed that participants were relatively insensitive to violations of the Newtonian principle of energy conservation. Together with the results of that study, the results of the current Experiment 1 are more consistent with the idea that people's intuitive understanding of physical events is based on heuristic processes (see also diSessa, 1993; Gilden & Proffitt, 1994; Hecht, 2001; McCloskey, 1983) than with the idea that it is based on the internalization of physical constraints (Sanborn et al., 2013).

Acknowledgments

I wish to thank Prof. Luigi Burigana for useful comments and suggestions.

References

Beasley, N. A. (1968). The extent of individual differences in the perception of causality. *Canadian Journal of Psychology, 22*, 399–407. http://dx.doi.org/10.1037/h0082779.

Boyle, D. G. (1960). A contribution to the study of phenomenal causation. *Quarterly Journal of Experimental Psychology, 12*, 171–179. http://dx.doi.org/10.1080/17470216008416721.

Brown, H. V., & Miles, T. R. (1969). Prior stimulation and the perception of causality. *Quarterly Journal of Experimental Psychology, 21*, 134–136. http://dx.doi.org/10.1080/14640746908400205.

Buehner, M. J., & Humphreys, G. R. (2010). Causal contraction: Spatial binding in the perception of collision events. *Psychological Science, 21*, 44–48. http://dx.doi.org/10.1177/0956797609354735.

Bullock, M., Gelman, R., & Baillargeon, R. (1982). The development of causal reasoning. In W. Friedman (Ed.). *The developmental psychology of time* (pp. 209–255). London: Academic Press.

Caggiano, V., Fleischer, F., Pomper, J. K., Giese, M. A., & Thier, P. (2016). Mirror neurons in monkey premotor area F5 show tuning for critical features of visual causality perception. *Current Biology, 26*, 3077–3082. http://dx.doi.org/10.1016/j.cub.2016.10.007.

Choi, H., & Scholl, B. J. (2006). Measuring causal perception: Connections to representational momentum? *Acta Psychologica, 123*, 91–111. http://dx.doi.org/10.1016/j.actpsy.2006.06.001.

Clement, J. (1982). Students' preconceptions in introductory mechanics. *American Journal of Physics, 50*, 66–71. http://dx.doi.org/10.1119//1.12989.

Firestone, C., & Scholl, B. J. (2016). Cognition does not affect perception: Evaluating the evidence for "top-down" effects. *Behavioral and Brain Sciences, 39*, e229. http://dx.doi.org/10.1017/S0140525X15000965.

Fonlupt, P. (2003). Perception and judgment of physical causality involve different brain structures. *Cognitive Brain Research, 17*, 248–254. http://dx.doi.org/10.1016/S0926-6410(03)00112-5.

Gemelli, A., & Cappellini, A. (1959). The influence of the subject's attitude in perception. *Acta Psychologica, 14*, 12–23. http://dx.doi.org/10.1016/0001-6918(58)90003-9.

Gilden, D. L., & Proffitt, D. R. (1994). Heuristic judgment of mass ratio in two-body collisions. *Perception & Psychophysics, 56*, 708–720. http://dx.doi.org/10.3758/BF03208364.

Gruber, H. E., Fink, C. D., & Damm, V. (1957). Effects of experience on perception of causality. *Journal of Experimental Psychology, 53*, 89–93. http://dx.doi.org/10.1037/h0048506.

Halloun, I. A., & Hestenes, D. (1985). Common sense concepts about motion. *American Journal of Physics, 53*, 1056–1065. http://dx.doi.org/10.1119/1.14031.

Hecht, H. (2001). Regularities of the physical world and the absence of their internalization. *Behavioral and Brain Sciences, 24*, 608–617. http://dx.doi.org/10.1017/S0140525X01000036.

Hubbard, T. L. (2013a). Phenomenal causality I: Varieties and variables. *Axiomathes, 23*, 1–42. http://dx.doi.org/10.1007/s10516-012-9198-8.

Hubbard, T. L. (2013b). Phenomenal causality II: Integration and implication. *Axiomathes, 23*, 485–524. http://dx.doi.org/10.1007/s10516-012-9200-5.

Hubbard, T. L. (2013c). Launching, entraining, and representational momentum: Evidence consistent with an impetus heuristic in perception of causality. *Axiomathes, 23*, 633–643. http://dx.doi.org/10.1007/s10516-012-9186-z.

Hubbard, T. L., & Ruppel, S. E. (2002). A possible role of naive impetus in Michotte's "launching effect:" evidence from representational momentum. *Visual Cognition, 9*, 153–176. http://dx.doi.org/10.1080/13506280143000377.

Hubbard, T. L., & Ruppel, S. E. (2013). Ratings of causality and force in launching and shattering. *Visual Cognition, 21*, 987–1009. http://dx.doi.org/10.1080/13506285.2013.847883.

Hume, D. (1977). *An enquiry concerning human understanding.* Indianapolis: Hackett (Original work published 1748).

Kaiser, M. K., & Proffitt, D. R. (1987). Observers' sensitivity to dynamic anomalies in collisions. *Perception & Psychophysics, 42*, 275–280. http://dx.doi.org/10.3758/BF03203079.

Kim, S.-H., Feldman, J., & Singh, M. (2013). Perceived causality can alter the perceived trajectory of apparent motion. *Psychological Science, 24*, 575–582. http://dx.doi.org/10.1177/0956797612458529.

Leslie, A. M., & Keeble, S. (1987). Do six-month-old infants perceive causality? *Cognition, 25*, 265–288. http://dx.doi.org/10.1016/S0010-0277(87)80006-9.

Mascalzoni, E., Regolin, L., & Vallortigara, G. (2010). Innate sensitivity for self-propelled causal agency in newly hatched chicks. *PNAS, 107*, 4483–4485. http://dx.doi.org/10.1073/pnas.0908792107.

Matsuno, T., & Tomonaga, M. (2017). Causal capture effects in chimpanzees (*Pan troglodytes*). *Cognition, 158*, 153–164. https://doi.org/10.1016/j.cognition.2016.10.023.

McCloskey, M. (1983). Intuitive physics. *Scientific American, 248*, 122–130. http://dx.doi.org/10.1038/scientificamerican0483-122.

Michotte, A. (1963). *The perception of causality.* London, UK: Methuen (Original work published 1946).

Moors, P., Wagemans, J., & de-Wit, L. (2017). Causal events enter awareness faster than non-causal events. *PeerJ, 5*, e2932. http://dx.doi.org/10.7717/peerj.2932.

Newman, G. E., Choi, H., Wynn, K., & Scholl, B. J. (2008). The origins of causal perception: Evidence from postdictive processing in infancy. *Cognitive Psychology, 57*, 262–291. http://dx.doi.org/10.1016/j.cogpsych.2008.02.003.

Piaget, J. (1970). Piaget's theory. In P. H. Mussen (Vol. Ed.), *Carmichael's manual of child psychology. Vol. 1. Carmichael's manual of child psychology* (pp. 703–732). New York: Wiley.

Powesland, P. F. (1959). The effect of practice upon the perception of causality. *Canadian Journal of Psychology, 13*, 155–168. http://dx.doi.org/10.1037/h0083773.

Rips, L. J. (2011). Causation from perception. *Perspectives on Psychological Science, 6*, 77–97. http://dx.doi.org/10.1177/1745691610393525.

Rolfs, M., Dambacher, M., & Cavanagh, P. (2013). Visual adaptation of the perception of causality. *Current Biology, 23*, 250–254. http://dx.doi.org/10.1016/j.cub.2012.12.017.

Roser, M. E., Fugelsang, J. A., Dunbar, K. N., Corballis, P. M., & Gazzaniga, M. S. (2005). Dissociating processes supporting causal perception and causal inference in the brain. *Neuropsychology, 19*, 591–602. http://dx.doi.org/10.1037/0894-4105.19.5.591.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review, 120*, 411–437. http://dx.doi.org/10.1037/a0031912.

Schlottmann, A. (1999). Seeing it happen and knowing how it works: How children understand the relationship between perceptual causality and underlying mechanism. *Developmental Psychology, 35*, 303–317. http://dx.doi.org/10.1037/0012-1649.35.1.303.

Schlottmann, A. (2000). Is perception of causality modular? *Trends in Cognitive Sciences, 4*, 441–442. http://dx.doi.org/10.1016/S1364-6613(00)01555-2.

Schlottmann, A. (2001). Perception versus knowledge of cause and effect in children: When seeing is believing. *Current Directions in Psychological Science, 10*, 111–115. http://dx.doi.org/10.1111/1467-8721.00128.

Schlottmann, A., & Anderson, N. H. (1993). An information integration approach to phenomenal causality. *Memory and Cognition, 21*, 785–801. http://dx.doi.org/10.3758/BF03202746.

Scholl, B. J., & Nakayama, K. (2002). Causal capture: Contextual effects on the perception of collision events. *Psychological Science, 13*, 493–498. http://dx.doi.org/10.1111/1467-9280.00487.

Scholl, B. J., & Nakayama, K. (2004). Illusory causal crescents: Misperceived spatial relations due to perceived causality. *Perception, 33*, 455–469. http://dx.doi.org/10.1068/p5172.

Scholl, B. J., & Tremoulet, P. D. (2000). Perceptual causality and animacy. *Trends in Cognitive Sciences, 4*, 299–309. http://dx.doi.org/10.1016/S1364-6613(00)01506-0.

Schultz, T. R., Fisher, G. W., Pratt, C. C., & Rulf, S. (1986). Selection of causal rules. *Child Development, 57*, 143–152. http://dx.doi.org/10.2307/1130646.

diSessa, A. A. (1993). Toward an epistemology of physics. *Cognition and Instruction, 10*, 105–225. http://dx.doi.org/10.1080/07370008.1985.9649008.

Straube, B., & Chatterjee, A. (2010). Space and time in perceptual causality. *Frontiers in Human Neuroscience, 4*. http://dx.doi.org/10.3389/fnhum.2010.00028 (Article 28).

Vicovaro, M., & Burigana, L. (2014). Intuitive understanding of the relation between velocities and masses in simulated collisions. *Visual Cognition, 22*, 896–919. http://dx.doi.org/10.1080/13506285.2014.933940.

Vicovaro, M., & Burigana, L. (2016). Intuitive understanding of the relationship between the elasticity of objects and kinematic patterns of collisions. *Attention, Perception, & Psychophysics, 78*, 618–635. http://dx.doi.org/10.3758/s13414-015-1033-z.

Wagemans, J., van Lier, R., & Scholl, B. J. (2006). Introduction to Michotte's heritage in perception and cognition research. *Acta Psychologica, 123*, 1–19. http://dx.doi.org/10.1016/j.actpsy.2006.06.003.

White, P. A. (2007). Impressions of force in visual perception of collision events: A test of the causal asymmetry hypothesis. *Psychonomic Bulletin & Review, 14*, 647–652. http://dx.doi.org/10.3758/BF03196815.

White, P. A. (2009a). Property transmission: An explanatory account of the role of similarity information in causal inference. *Psychological Bulletin, 135*, 774–793. http://dx.doi.org/10.1037/a0016970.

White, P. A. (2009b). Perception of forces exerted by objects in collision events. *Psychological Review, 116*, 580–601. http://dx.doi.org/10.1037/a0016337.

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. http://dx.doi.org/10.2307/1423411.

White, P. A., & Milne, A. (1999). Impressions of enforced disintegration and bursting in the visual perception of collision events. *Journal of Experimental Psychology: General, 128*, 499–516. http://dx.doi.org/10.1037/0096-3445.128.4.499.

White, P. A., & Milne, A. (2003). Visual impressions of penetration in the perception of objects in motion. *Visual Cognition, 10*, 605–619. http://dx.doi.org/10.1080/13506280344000013.

Wolff, P. (2007). Representing causation. *Journal of Experimental Psychology: General, 136*, 82–111. http://dx.doi.org/10.1037/0096-3445.136.1.82.

Young, M. E., & Falmier, O. (2008). Color change as a causal agent revisited. *American Journal of Psychology, 121*, 129–156. http://dx.doi.org/10.2307/20445447.