Investigating causal inference difficulties with a simple, qualitative force-and-motion problem

Sara Jaramillo
Department of Psychology, University of Pittsburgh, 3420 Forbes Ave., Pittsburgh, PA, 15260

Eric Kuo
Departments of Physics and Curriculum & Instruction, University of Illinois at Urbana-Champaign, 1110 W. Green St., Urbana, IL, 61801

Benjamin M. Rottman and Timothy J. Nokes-Malach
Department of Psychology, University of Pittsburgh, 3420 Forbes Ave., Pittsburgh, PA, 15260

Understanding the nature of causality is a key component of conceptual understanding in science. The hypothesis of this study is that certain types of causal inference are more challenging than others. If correct, particularly challenging causal inferences may provide a unified explanation for different conceptual difficulties across physics content areas. This paper investigates two aspects of a simple, qualitative force-and-motion problem that may impact the difficulty of the causal inferences required. The first aspect is the type of causal inference required: cause-to-effect (CE), effect-to-cause (EC), or cause-to-cause (CC). The second aspect is information about an alternative cause, which can be explicitly constant, explicitly unknown, or ambiguous. To test the impact of these two problem aspects on participants inference accuracy we conducted an on-line experiment in which participants were randomly assigned to one of thirty-six conditions that systematically varied these two aspects across conditions. The results show that (i) for explicitly constant alternative causes, CC inferences are more difficult than CE or EC inferences, (ii) inferences given explicitly unknown alternative cause information are more difficult than inferences given explicitly constant alternative causes, and (iii) ambiguous alternative cause information is treated as implying the alternative causes are explicitly constant, which is in line with conversational assumptions rather than a formal, logical perspective. These results hint at the potential fruitfulness of understanding the causal inferences underlying conceptual difficulties in physics.

I. INTRODUCTION

Understanding the nature of causality is a key component of conceptual understanding in science [1–4]. diSessa and Sherin [1] proposed that part of conceptual understanding is knowledge that allows one to draw causal inferences from information read out from physical situations. We investigated the hypothesis of whether certain types of causal inference are more challenging than others. If correct, particularly challenging causal inferences may provide a unified explanation for different conceptual difficulties across physics content areas.

A physical system's causal structure determines the types of causal inferences that can be made. The simplest, interesting causal structure is a common-effect structure where two independent causes impact one effect. Consider this experimental situation: a cart with an attached sail is placed in a wind tunnel where the fan blows the wind, the force of the wind acts on the sail, and the cart accelerates. The causal diagram representing this system (Fig. 1) indicates that the force of the fan and the mass of the cart are two independent *causes* that impact an *effect*, the acceleration of the cart. Fig. 1 also indicates that each cause is linked to a *source object*, either the fan used or the cart used. This system can be altered by intervening on the source objects (e.g., the fan can be changed to increase the force of the wind on the sail). Although there are other potential source objects that could be altered to affect the causal factors, here we work with a simplified system. For this system, there are three types of causal inferences that can be made from changes in the system (Table I):

- cause-to-effect (CE) inference: a change in a cause (force or mass) is used to infer whether there is a change in the effect (acceleration).
- effect-to-cause (EC) inference: a change in the effect (acceleration) is used to infer whether there is a change in a cause (force or mass).
- cause-to-cause (CC) inference: a change in one cause is used to infer whether there is a change in another cause.

For independent causes, changes in one cause will not affect the other cause and recognizing this independence between causes is part of physical conceptual understanding.

Errors in making correct cause-cause inferences may describe a subset of conceptual difficulties seen in physics education. The classic Newton's 3rd law error of inferring that, in a collision between two objects, a less massive object will feel a larger force [5,6] can be interpreted as an incorrect cause-cause inference: using information about the difference in masses to infer a difference in forces. Redish and Kuo [7] discussed an introductory E&M problem where a test charge in an external electric field feels a force. Students commonly answered incorrectly that changing the test charge would change the external electric field, even
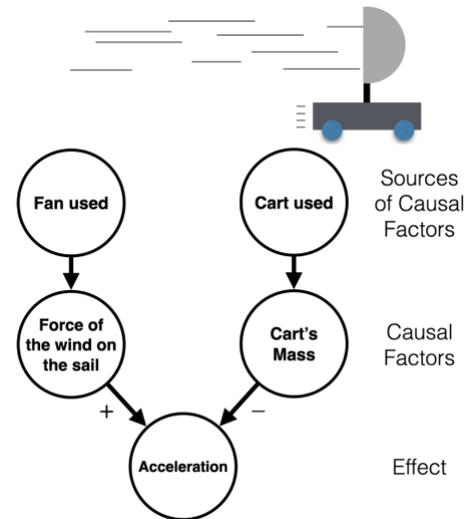

FIG. 1. The causal structure of the sail-cart situation. The + sign indicates that if the force increases, acceleration will increase. The – sign indicates that if the mass increases, acceleration will decrease.

though the test charge and the external electric field are independent causes that impact the force on the test charge. In a previous study [8] a question about a spring attached to the ceiling being stretched out by a hanging block was used. A common error was inferring that increasing the mass of the block changed the "spring stiffness," even though the weight of the block and spring constant are independent causes that impact the distance stretched. Motivated by these examples, the current study focuses on CC inferences, testing if these are easier or more difficult than CE and EC inferences.

Some conceptual difficulties in physics can also be modeled as the ability to recognize and integrate the presence of alternative causes in causal inference [9,10]. In a CE or EC inference, information about an alternative cause can affect the normative inferences that can be made. For instance, if the fan is replaced with another fan to increase the force of the wind, the acceleration may not increase, since mass could increase to offset or overcome the impact of the changing force. Similarly, if the acceleration increases, one may not be able to infer that the force of the fan increases since the change to acceleration may be due to a decrease in the mass. Therefore, this study also investigates the impact of different alternative cause information on the difficulty of causal inferences.

TABLE I. Structure of the three causal inference types.

Inference type	What changes?	Inferred change
Cause-Effect (CE)	Force / Mass	Acceleration
Effect-Cause (EC)	Acceleration	Force/Mass
Cause-Cause (CC)	Force/Mass	Mass/Force

II. RESEARCH QUESTIONS

This paper investigates two aspects of qualitative physics problems that may impact the difficulty of the causal inferences required. The first is the type of causal inference asked: CE, EC, or CC. Here, we focus on whether CC inferences are easier or more difficult than CE and EC inferences, since, as discussed earlier, they may describe many known conceptual difficulties in physics.

The second aspect under investigation is alternative cause information. This study includes causal inferences when the alternative source that determines the alternative cause is explicitly held constant or is explicitly allowed to vary. The question is whether one of these alternative source conditions is more difficult than the other. We also investigate two competing hypotheses of how people make causal inferences when information about the alternative source object is left ambiguous. One hypothesis follows from the pragmatic, conversational assumption that all relevant information is provided [11]. From this conversational pragmatics perspective, the absence of explicit information about an alternative source could lead people to assume—implicitly or explicitly—that this alternative cause is held constant. A second hypothesis follows a more formal logic perspective: an unmentioned alternative source could vary, leading one to the assumption that the alternative cause status is unknown, so there is no way to make a definite causal inference. This difference in perspectives is perhaps why students can view some questions as "trick questions," as students may make pragmatic conversational assumptions while they are graded from a logical perspective.

III. METHODS

The data collection and analysis plan were preregistered on the Open Science Framework [https://osf.io/95g72/].

Participants were recruited online from Amazon Mechanical Turk and were all located in the US. Data were collected until 900 participants successfully completed the study and the embedded attention and verification checks. Two participants who failed to complete the study were included in data collection by error. They were excluded from the analysis, leaving $N = 898$ participants.

The experimental design was a 6 (causal question) x 3 (alternative source information) between-subjects design, with ~50 participants in each condition. Each participant answered one inference question about the sail-cart. We label the causes and effects using subscripts—for example, E_aC_m is the EC question where information about the acceleration (a) is used to draw an inference about the mass (m). There are 6 causal question permutations: two CE questions (C_FE_a, C_mE_a), two EC questions (E_aC_F, E_aC_m), and two CC questions (C_FC_m, C_mC_F). There are also 3 alternative source information types: Explicitly Constant, Explicitly Unknown, or Ambiguous.

Participants first read a brief passage about how force (F) and mass (m) can impact an object's acceleration (a). Then, participants were presented with a sail-cart question. Table II provides the full question text for one CE, EC, and CC question, with three possible alternative source information statements. Each question included:

- An introduction laying out the physical situation of a sail-cart accelerating in a wind tunnel.
- A change to the system—an increase to either the force, mass, or acceleration.
- Alternative source information in either the explicitly constant, explicitly varying, or ambiguous condition.
- A question asking participants to infer how another factor of the system changes: is it greater, the same, less, or indeterminate?

TABLE II. Problem text presented for one of the CE, EC, and CC questions.

Section of Problem Text		Causal Inference Type		
		C_FE_a	E_aC_m	C_FC_m
	Introduction	*A small cart with a sail is placed in a wind tunnel. A fan in the wind tunnel spins at a constant rate, generating the wind. The wind hits the sail and the sail-cart starts accelerating at a certain rate.*		
	Change to System	*In a second trial, the fan was replaced with another fan that exerts 3 times more force than before.*	*In a second trial, the sail-cart's acceleration was 3 times greater than before.*	*In a second trial, the fan was replaced with another fan that exerts 3 times more force than before.*
1 of 3 Alternative Source Information Options	Explicitly Constant	*The cart was kept the same as before.*	*The fan was kept the same as before.*	*The cart was kept the same as before.*
	Explicitly Unknown	*It's unknown if the cart changed or was kept the same as before.*	*It's unknown if the fan changed or was kept the same as before.*	*It's unknown if the cart changed or was kept the same as before.*
	Ambiguous	*None*	*None*	*None*
	Inference Question	*In this case, the acceleration of the sail-cart would be?*	*In this case, the mass of the cart would be?*	*In this case, the mass of the cart would be?*
	Response Choices	*Greater than before; Less than before; The same as before; It cannot be determined from the information given.*		

A. Coding responses

Table III shows the normative responses given Explicitly Constant and Explicitly Unknown alternative source information. For the Explicitly Constant condition, the normative response corresponds to the direct connection between factors, as shown in Figure 1. In the Explicitly Unknown condition, the normative response is always "it cannot be determined" (denoted with ?), because when it is unknown whether the alternative source changes, definite qualitative inferences about a target factor cannot be made.

For example, the C_FE_A question states that the fan is replaced such that the force of the wind increases and asks about the acceleration. In the Explicitly Constant condition, the cart stays the same, so the mass is held constant, and an increase to force means the acceleration will be greater than before (>). In the Explicitly Unknown condition, it is unknown whether the cart stays the same or changes, so the mass can vary, and it cannot be determined what happens to acceleration (?). The other responses are less than before (<) or the same as before (=).

For each of the 18 different conditions, responses were coded into these three categories: normative for explicitly constant, Normative for explicitly unknown, or other responses. For the Explicitly Constant and Explicitly Unknown conditions, we are interested in how frequently participants provide the normative response for the question given. For the Ambiguous condition, conversational or logical assumptions both be appropriate depending on context, so we do not label one response as being normative. Rather, we are interested in whether the distribution of responses aligns more with that for the Explicitly Constant condition (in line with conversational assumptions) or the Explicitly Unknown condition (in line with the formal logic perspective).

IV. RESULTS

We aggregated our results from 6 causal questions into the 3 causal inference types: CE, EC, and CC. Although there were some differences between questions within causal inference type, they were not germane to the overall results here.

TABLE III. Response coding. Key: greater than before (>), less than before (<), the same as before (=), it cannot be determined (?)

Causal Question	Normative Response for Explicitly Constant	Normative Response for Explicitly Unknown	Other Responses
C_FE_a	>	?	<, =
C_mE_a	<	?	>, =
E_aC_F	>	?	<, =
E_aC_m	<	?	>, =
C_FC_m	=	?	>, <
C_mC_F	=	?	>, <

A. Does the type of causal inference or explicitly constant / explicitly varying alternative source information impact the difficulty of causal inferences?

Figure 2 shows the proportion of normative responses for the Explicitly Constant and Explicitly Unknown conditions. We performed a 3 (causal inference type: CE, EC, CC) x 2 (alternative source information: explicitly constant, explicitly unknown) logistic regression using log odds-ratio of normative responses as the dependent variable. CC-explicitly constant was chosen as the reference group because of our focus on whether CC inferences are easier or more difficult CE or EC inference.
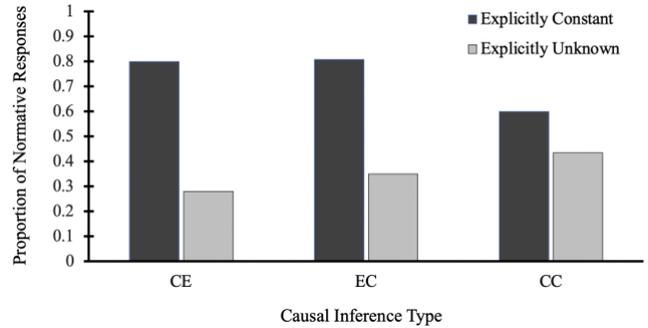

FIG. 2. Proportion of normative responses for CE, EC, CC causal inferences in the Explicitly Constant and Explicitly Unknown conditions.

Table IV shows the results of the logistic regression model. Odds ratios greater than one indicate more normative responses (and odds ratios less than one indicate fewer normative responses). The significant effects of CE and EC indicate that, in the Explicitly Constant condition, CE and EC inferences had more normative responses than CC inferences. CC inferences have greater rates of "other" responses (as shown in Fig. 3)—such as, increasing the force of the fan changes the mass of the cart, even though the cart used is explicitly stated to be kept the same.

The significant effect of Explicitly Unknown indicates that, for CC inferences, Explicitly Unknown questions had fewer normative responses than Explicitly Constant questions. This appears to be due to responses on explicitly unknown questions that say that a cause stays the same, even

TABLE IV. Results of 3 (causal inference type) x 2 (explicitly constant, explicitly unknown) logistic regression with log odds of normative responses as the dependent variable.

Effect	b	Odds Ratio [95% CI]	p
CE	0.98	2.67 [1.42, 5.02]	0.002
EC	1.03	2.81 [1.48, 5.33]	0.002
Explicitly Unknown	-0.67	0.51 [0.29, 0.90]	0.019
CE x Explicitly Unknown	-1.66	0.19 [0.08, 0.45]	<0.001
EC x Explicitly Unknown	-1.39	0.25 [0.11, 0.59]	0.002

though it is explicitly unknown if its source object changes (as shown in Fig. 3)—such as, responding that the mass of the cart stays the same even though it is unknown whether the cart used changes or stays the same.

The significant interactions of CE x Explicitly Unknown and EC x Explicitly Unknown indicate that the difference in normative responses between explicitly constant and explicitly unknown is greater for CE and EC inferences than for CC inferences.

B. Is ambiguous alternative source information treated as constant or unknown?

Figure 3 shows the distribution of responses for ambiguous questions as compared to explicitly constant and explicitly unknown ones. By inspection, the distribution of responses for ambiguous CE, EC, and CC inference questions resembles the distributions for explicitly constant questions instead of explicitly unknown questions. To test this, we conducted a series of Fisher's exact tests for CC, CE, and EC separately, to see if the response distributions are different for the Ambiguous condition vs. the Explicitly Constant condition, and also to see if the response distributions are different for the Ambiguous condition vs. the Explicitly Unknown condition.

Response distributions to ambiguous questions are not different than distributions for explicitly constant questions for CE inferences (Fisher's exact $p = 0.91$), EC inferences (Fisher's exact $p = 0.23$), and CC inferences (Fisher's exact $p = 0.37$). By contrast, ambiguous question response distributions are different than those for explicitly unknown
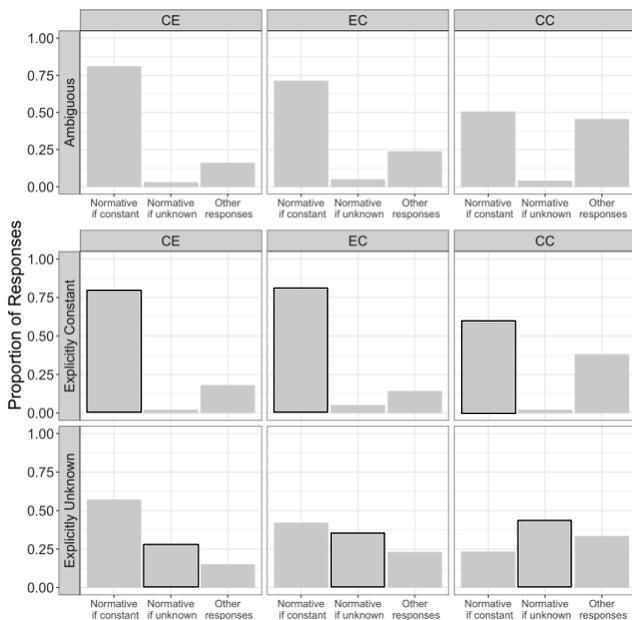

FIG. 3. Response distributions for each causal inference type across different alternative source conditions. The outlined bar for explicitly constant and explicitly unknown conditions indicates the normative response.

questions for CE, EC, and CC inferences (all Fisher's exact p's < 0.001).

V. DISCUSSION

In the current study, we investigated how different aspects of causal inference questions impact how easy or difficult normative inferences are. One aspect was the type of causal inference asked: CE, EC, or CC. For explicitly constant alternative source information, we found that CC inferences here are more difficult than CE or EC inferences. It appears that people sometimes believe that changing one independent cause can impact another.

Another aspect of causal inference questions was alternative source information. Participants had more difficulty answering explicitly unknown alternative source questions than the explicitly constant questions. The relatively high rates of "normative if constant" responses on explicitly unknown questions is consistent with other research showing that people often fail to integrate potentially varying alternative causes into their causal inferences, on both physics problems [9,10] and in other contexts [12–14]. Why might this be? It is possible that this statement is ignored or that people did not know what to do with this information. As a part of this study, we have also collected participants' open-response explanations of their reasoning and asked them to explicitly state their assumptions about the alternative source. In future work, we will use these responses to evaluate possible explanations.

Further, we found that ambiguous alternative sources were treated as if they were explicitly constant, consistent with the conversational pragmatics hypothesis. In future analyses we will also use the open-response explanations to differentiate whether participants make an explicit assumption that unmentioned alternative causes are held constant or whether they neglect those unmentioned alternative causes.

The overall premise of this program of research is that systematically conceptualizing physics questions according to causal inference type, and understanding the causal and pragmatic reasoning used on these questions, may provide a unified explanation for conceptual difficulties across physics content areas. Future work can investigate whether the pattern of findings here would be found for other physical topics, situations, and questions. If difficult causal inferences do underlie a subset of conceptual difficulties, developing physics instruction around causality, causal inference, and pragmatic assumptions may improve conceptual understanding across different content areas.

[1] A. A. diSessa and B. L. Sherin, *What Changes in Conceptual Change?*, International Journal of Science Education **20**, 1155 (1998).

[2] B. Y. White, *ThinkerTools: Causal Models, Conceptual Change, and Science Education*, Cognition and Instruction **10**, 1 (1993).

[3] S. Carey, *Sources of Conceptual Change*, in *Conceptual Development: Piaget's Legacy*, edited by E. K. Scholnick, K. Nelson, S. A. Gelman, and P. H. Miller (Lawrence Erlbaum, Mahwah, NJ, 1999), pp. 293–326.

[4] M. Braaten and M. Windschitl, *Working toward a Stronger Conceptualization of Scientific Explanation for Science Education*, Science Education **95**, 639 (2011).

[5] A. Elby, *Helping Physics Students Learn How to Learn*, American Journal of Physics, Physics Education Research Supplement **69**, S54 (2001).

[6] D. E. Brown, *Students' Concept of Force: The Importance of Understanding Newton's Third Law*, Physics Education **24**, 353 (1989).

[7] E. F. Redish and E. Kuo, *Language of Physics, Language of Math: Disciplinary Culture and Dynamic Epistemology*, Sci & Educ **24**, 561 (2015).

[8] E. Kuo, B. M. Rottman, and T. J. Nokes-Malach, T.J. *Investigating Causal Inference in Physics to Understand Student Difficulties*, presented at the summer meeting of the American Association of Physics Teachers, online, 2020 (unpublished).

[9] E. Kuo, N. K. Weinlader, B. M. Rottman, and T. J. Nokes-Malach. *Using causal networks to examine resource productivity and coordination in learning science*, in The Interdisciplinary of the Learning Sciences, 14th International Conference of the Learning Sciences (ICLS) 2020, Volume 2, Nashville, Tennessee, edited by M. Gresalfi and I. S. Horn (International Society of the Learning Sciences, 2020), pp. 875-876.

[10] A. F. Heckler and A. M. Bogdan, *Reasoning with Alternative Explanations in Physics: The Cognitive Accessibility Rule*, Physical Review Physics Education Research **14**, 010120 (2018).

[11] H. P. Grice, *Logic and Conversation*, in *Speech Acts* (Brill, 1975), pp. 41–58.

[12] P. M. Fernbach and B. Rehder, *Cognitive Shortcuts in Causal Inference*, Argument & Computation **4**, 64 (2013).

[13] P. M. Fernbach, A. Darlow, and S. A. Sloman, *Asymmetries in Predictive and Diagnostic Reasoning.*, Journal of Experimental Psychology: General **140**, 168 (2011).

[14] P. M. Fernbach, A. Darlow, and S. A. Sloman, *Neglect of Alternative Causes in Predictive but Not Diagnostic Reasoning*, Psychological Science **21**, 329 (2010).