THINKING & REASONING, 2017
VOL. 23, NO. 3, 292–317
https://doi.org/10.1080/13546783.2017.1316314

Routledge
Taylor & Francis Group

Check for updates

Dual frames for causal induction: the normative and the heuristic

Ikuko Hattori [a], Masasi Hattori [b], David E. Over [c], Tatsuji Takahashi [d] and Jean Baratgin [e]

[a]College of Letters, Ritsumeikan University, Kyoto, Japan; [b]College of Comprehensive Psychology, Ritsumeikan University, Osaka, Japan; [c]Psychology Department, Durham University, Durham, UK; [d]School of Science and Engineering, Tokyo Denki University, Saitama, Japan; [e]Laboratoire CHArt, University of Paris 8 and Institut Jean Nicod, Paris, France

ABSTRACT
Causal induction in the real world often has to be quick and efficient as well as accurate. We propose that people use two different frames to achieve these goals. The A-frame consists of heuristic processes that presuppose rarity and can detect causally relevant factors quickly. The B-frame consists of analytic processes that can be highly accurate in detecting actual causes. Our dual frame theory implies that several factors affect whether people use the A-frame or the B-frame in causal induction: among these are symmetrical negation, intervention and commitment. This theory is tested and sustained in two experiments. The results also provide broad support for dual process accounts of human thinking in general.

ARTICLE HISTORY Received 28 May 2015; Accepted 30 March 2017

KEYWORDS Causal induction; dual frames; negation; intervention; commitment

Introduction

Causal induction is vital for surviving and prospering in an uncertain world. It enables people to predict and control their future. But how do people infer inductively that one event is the cause of another? They might quickly come to one conclusion about causation, but revise it over time and arrive at quite a different conclusion in the end. For an example, suppose some friends become unwell after eating sushi at a new restaurant. We might immediately form the belief that eating the sushi caused our friends illness. As the days pass, however, this belief may weaken. We may hear of other friends who ate similar sushi at the same restaurant and did not become unwell, and still others who become unwell in the same way without eating the sushi. We would give up the belief completely if we read in a newspaper that the health

CONTACT Ikuko Hattori iku.hattori@gmail.com; iht23083@pl.ritsumei.ac.jp
Supplemental data for this article can be accessed at http://dx.doi.org/10.1080/13546783.2017.1316314.

Table 1. A 2 × 2 contingency table representing covariation information between a candidate cause (*C*) and a target effect (*E*).

	E	¬*E*
C	*a*	*b*
¬*C*	*c*	*d*

Note: *C* and ¬*C* represent the occurrence and non-occurrence of a candidate cause, and *E* and ¬*E* represent the occurrence and non-occurrence of a target effect, respectively.

authorities have investigated the restaurant and have found nothing wrong with the sushi after testing it.

There are four possible events that are clearly relevant, in some way, to such questions about causation. People can observe the sushi eating and illness, or the sushi eating and no illness. They can also observe no sushi eating and illness, or no sushi eating and no illness. More formally, let *C* be a proposed cause of an effect *E*. There are the four Boolean possibilities: *C & E*, *C & not-E*, *not-C & E* and *not-C & not-E*. There can be observations of each of these cases, and these can be collected in a 2 × 2 contingency table (see Table 1), in which the *a*-cell has the number of *C & E* observations, the *b*-cell the number of *C & not-E* observations, the *c*-cell the number of *not-C & E* observations and the *d*-cell the number of *not-C & not-E* observations. People could be lucky and find that there are only *a*-cell cases and *d*-cells cases. Perhaps everyone who ate the sushi at the restaurant became unwell, and no one become unwell who ate something else at the restaurant. It would then be easy to infer that the sushi caused the illness.

More generally, for causal claims, all the cells, *a*, *b*, *c* and *d*, could have at least some entries. In this much more common type of example, the most cited normative proposal is to use the Δ*P* relation (e.g., Anderson & Sheu, 1995; Cheng, 1997; Jenkins & Ward, 1965; Kao & Wasserman, 1993; Waldmann, 1996).[1] There are good grounds for inferring that *C* causes *E* when the probability of *E* given *C*, $P(E|C)$, is significantly higher than the probability of *E* given *not-C*, $P(E|\neg C)$:

$$\Delta P = P(E|C) - P(E|\neg C) = \frac{a}{a+b} - \frac{c}{c+d} \quad (1)$$

According to Jenkins and Ward (1965), Δ*P* is an index of the actual degree of control or contingency between the action and the outcomes, rather than a dependency or correlation, and the magnitude of Δ*P* is held to provide an index of the amount of control. In the example above, sushi eating at the restaurant could turn out to be poorly justified by Δ*P*. There

[1]This relation is normative because it corresponds also to a Bayesian measure of confirmation. There are some other measures in the literature (Fitelson & Hitchcock, 2011).

might be many cases of sushi eating at the place and no illness, and so entries in the *b*-cell, and there could also be some entries in the *c*-cell and many in the *d*-cell.

Although ΔP has a rational justification, its descriptive adequacy has been questioned, in particular because people tend to disregard *d*-cell events (e.g., Kao & Wasserman, 1993; McKenzie & Mikkelsen, 2007; Schustack & Sternberg, 1981). In response, Hattori and Oaksford (2007) proposed the *dual factor heuristic* model, *DFH*, as an alternative. They define DFH in the following way:

$$\mathrm{DFH} = \sqrt{P(E|C)\ P(C|E)} = \frac{a}{\sqrt{(a+b)(a+c)}} \tag{2}$$

DFH is an index for the degree of relevance between two events as a candidate cause and an effect. According to Hattori and Oaksford (2007), causal induction consists of two stages: a *heuristic stage* for distinguishing relevant causal candidates from irrelevant factors, and an *analytic stage* for discriminating between genuine and spurious causations. In the heuristic stage of causal induction, it is argued that people do not strictly distinguish between causation and correlation, and they prioritise speed rather than accuracy. Thus DFH is defined as the limit of correlation coefficient φ for a 2×2 contingency table (i.e., a counterpart of Pearson's correlation coefficient for two continuous variables) when the *d*-cell goes to infinity: $\lim_{d \to \infty} \varphi$. This position is motivated by noting the usual relative *rarity* of events in causal claims (Hattori, 2002; McKenzie & Mikkelsen, 2000, 2007; Oaksford & Chater, 1994, 2003). The candidate cause-and-effect events *C* and *E* tend to be "rare", with relatively small set sizes, while the *not-C* events and *not-E* events tend to be common, with relatively large set sizes.

Consider another question about causation, "Does eating fly agaric mushrooms cause hallucinations?" Fly agaric mushroom eating is rare: few people do this eating. It is consequently efficient to look closely at these people to find out whether hallucinations follow, giving us information for the *a*- and *b*-cells of the 2×2 table. It would be grossly inefficient to examine people who do not eat fly agaric mushrooms, a gigantic set of people, to get information for the *c*- and *d*-cells. We already know that the vast majority of people do not eat these mushrooms and do not have hallucinations. The *d*-cell generally records the frequency with which nothing happens, the *not-C & not-E* cases. These cases are common when the occurrences of the causes and the effects are rare. Disregarding *d*-cell events takes a great burden off working memory. In this regard, DFH can be an effective heuristic for detecting causally relevant factors in the environment.

We hypothesise that causal induction has two aspects, characterised by a fast process that is mainly devoted to relevance detection, and by a slow process for inferring genuine causes. Relevance detection is described by DFH, while inferring actual causation is based on ΔP.[2] To establish causation, and not mere correlation, *intervention* in a system is essential, for it can screen out other possible factors or structures (see Pearl, 2000; Spirtes, Glymour, & Scheines, 1993). We see ΔP as an index of control, and control through intervention is fundamental to scientific experimentation. Perhaps people who eat fly agaric mushrooms also tend to eat other kinds of wild mushrooms. Then the possibility of a spurious correlation between fly agaric eating and hallucinations could be eliminated by intervening to make sure that a sample of volunteers only ate fly agaric mushrooms on some occasion.

To develop the example more, suppose we are investigating the cause of some hallucinations in a group of people who gathered wild mushrooms to eat. In a heuristic process, we could ask each person who had the hallucinations (H) which mushrooms they ate. Perhaps the fly agaric (F) would be the most common answer, making the probability of eating these mushrooms given hallucinations, $P(F|H)$, high. We might also find out that a high proportion of people who ate the fly agaric had hallucinations, making $P(H|F)$ high. These findings would focus our attention on the fly agaric, supporting the implication of DFH, given by Equation (2), that causal strength increases when both $P(E|C)$ and $P(C|E)$ are high. At this point, we could intervene, in an analytic process, by asking some volunteers to eat the fly agaric and nothing else.

Dual frames for causal induction

In accord with the claim that causal induction has two different stages, the heuristic and analytic (Hattori & Oaksford, 2007), we hypothesise that there are two different frames for the corresponding inferences, based on a new theoretical framework for thinking: dual frame theory (Hattori, 2014; Hattori, Over, Hattori, Takahashi, & Baratgin, 2016). We will call these the *A*- and *B-frames* for inferring the relation between two events, depending on context or purpose. Here “A” stands for “attentional” and “B” stands for “balanced”: see Table 2. Basically, the A-frame is used for heuristic judgements, and the B-frame for analytic judgements.

In the heuristic A-frame, people focus on positive events. Actual occurrences, C and E cases, grab the attention, and non-occurrences, *not-C* and *not-E* cases, are ignored. Humans and other animals often pay less attention to contexts in which the possible causal events do not happen. Pigeons, for example, can easily find that pecking a marked key results in a reward, but

[2]Although there are more recent rational models of causal induction other than ΔP, including Power PC (Cheng, 1997), Causal Support (Griffiths & Tenenbaum, 2005) and the SS power model (Lu et al., 2008), we do not address them. This is not only because ΔP is at the core of all the above rational models (e.g., ΔP is the numerator of the Power PC index), but also because our purpose is to contrast two frames, A and B, introduced later, of which, in our view, DFH and ΔP are representative, respectively.

Table 2. Characteristics of the A- and B-frames in causal reasoning.

	A-frame	B-frame
Epistemic aim	Fast screening	Control
Thought style	Relevance mode	Differentiation mode
Focalisation	Positivity focus	Comparative view
Psychological symmetry	Asymmetrical	Symmetrical
Negation	Explicit negation (X vs. $\neg X$)	Implicit negation (X vs. Y)
Cognitive process	Heuristic	Analytic
Cognitive resource	Low effort	High effort
D-cell	Disregard	Respect
Base rate	Rare	Not rare (moderate)
Causality scaling	Monopolar (null/effective)	Bipolar (preventive/generative)
Invasiveness	Observation	Intervention
Commitment	Uncommitted (low commitment)	Committed (high commitment)
Activeness	Passive	Active

they find it hard to learn that pecking a non-marked key results in a reward (Jenkins & Sainsbury, 1969; Jenkins & Sainsbury, 1970). The occurrence of an event and its non-occurrence are logical negations of each other, but psychologically speaking, this symmetry often does not exist. This psychological asymmetry leads to the ordinary attitude toward *d*-cell information in causal reasoning. People can readily notice the occurrence of an event, but are prone to ignore its non-occurrence, which is often the "default" status for them. They would be struck by the engine of a car not starting when the ignition key was turned on, but not by the engine not starting when the ignition key was not turned on. Disregarding the *d*-cell is reasonable, as noted before, when the occurrences of the candidate causal events are rare. Hattori (2014) pointed out that this psychological asymmetry between occurrence and non-occurrence is similar to the *figure–ground* relationship proposed by Rubin (1915/1958, 1921) as a purely perceptual phenomenon (see for detailed discussion, Hattori et al., 2016). In the A-frame, causal events are spotlighted "figures", and other events are in the vague background. The positivity focus of the heuristic A-frame allows us to make speedy and efficient causal judgements. After eating sushi at the new restaurant with friends, and hearing that they are unwell, we might predict that we will become unwell and make an appointment with a doctor immediately.

In contrast, in the analytic B-frame, people do not treat positive and negative events differently. They focus equally on the occurrence and non-occurrence of the event and make a comparison. The B-frame takes the place of A-frame when a more precise judgement is needed. Although correlation can be detected efficiently in the A-frame, reliably identifying precise causes, and distinguishing these from mere correlations, requires the B-frame. Causation is useful not only to predict the future, but also to control the environment. The degree of control can be defined, as ΔP describes, by the difference between, for instance, the probability of feeling unwell after eating at the new restaurant and that of feeling unwell after not eating there. Using

ΔP in the discovery that low hygiene standards at the new restaurant are the cause of the illness, the health authorities can close it down until improvements are made. Knowing precise causal relationships can be important, of course, and people can display an ability to do this. However, acquiring this precision can be costly, in time and cognitive load, and correlation alone can be useful in prediction and some decision-making. We hypothesise that people would use the A-frame as a default mode of thinking generally, but that they would try to switch the frame from A to B when B is less costly or of greater benefit in a given context.

In this research, we focus on three factors that can affect the change from one of these frames to the other in causal induction: *symmetrical negation*, *intervention* and *commitment*.

Consider *negation* first. X and $\neg X$ are logically related. The latter is the logical negation of the former, and they are complementary to each other. In our view, however, a statement X and its logical negation, $\neg X$, are not psychologically symmetrical. Compare "They ate the sushi" with "They did not eat the sushi." This is an example of the figure–ground relationship that we referred to before. We call this type of relation *non-symmetrical negation.* Non-symmetrical negation is compatible with the A-frame of thought. To search for the cause of feeling unwell, people will generally find it easier to focus on, or conceived of, sushi eating rather than of the non-action of not eating sushi, at least at the first stage of searching. In contrast, there are cases in which sushi eating (X) and, say, tempura eating (Y) are alternatives. There might be an occasion when everyone ate one or the other but not both, and therefore eating tempura (Y) means not-eating sushi ($\neg X$). We call this type of relation *symmetrical negation.* On these occasions, people will be more likely to pay attention to both sushi eating and non-sushi eating, i.e., tempura eating. The B-frame, with its comparative aspect, is likely to be activated when negation is symmetrical, simply because it is relatively easy to contrast specific alternatives, sushi eating and tempura eating in our example. Both X and its alternative Y, which equals $\neg X$ in logic, will then be focused on and compared.

Intervening on a system instead of observing it can also cause a shift of frames from A to B in causal induction, as suggested by Hattori and Oaksford (2007). In their Experiment 1, participants directed more attention to *d*-cell information when they decided whether to bring about an event rather than just passively observing it, perhaps because they switched to the B-frame. Compare deciding for oneself to use organic fertiliser on plants with merely observing the effect of organic fertiliser in general use. Hattori and Oaksford, however, did not control the cell frequency, and they did not evaluate their results using ΔP. We, therefore, examined this factor further.

The final factor is *commitment*. We also predict that this can cause a frame shift from A to B. Whether or not amateur gardeners use organic fertiliser on their flowers might not have serious consequences for them, and we say that they are not highly committed to the use of this fertiliser. In contrast, whether or not farmers use organic fertiliser on their main cash crop could have extremely serious economic consequences for them, making them highly committed, in our terms, to finding out whether this fertiliser is cost effective. The B-frame is more highly demanding of cognitive resources than the A-frame, but if the degree of commitment exceeds a certain point, perhaps in a cost–benefit trade-off, the B-frame can override the default A-frame.

We conducted two experiments to investigate how these three factors – symmetrical negation, intervention and commitment – are related to the two frames, A and B.

Experiment 1

In Experiment 1, the effects of symmetrical negation (non-symmetry vs. symmetry) and intervention (observation vs. intervention) on switching the frames were examined. We expected that participants' causal judgements would accord with ΔP, rather than DFH, in the symmetry–intervention condition, and that they would accord with DFH rather than ΔP in the non-symmetry–observation condition. According to dual frame theory, both symmetrical negation and intervention promote frame shifting from A to B. Under the B-frame, people pay attention, not only to positive events, but also to negative ones: they consider the *d*-cell as well as other cells. Therefore, ΔP would fit participants' causal judgements in the symmetry–intervention condition. On the other hand, DFH would fit well in the non-symmetry–observation condition, since most participants would stick to the default A-frame, where they focus on positive events and are prone to disregard negative events. The symmetry–observation and non-symmetry–intervention conditions have only one factor, either symmetrical negation or intervention. In the symmetry–observation condition, interaction of the factors might occur. That is, both models may fit to a certain degree with the judgements in this condition. Some participants may be able to switch to the demanding B-frame, but others may find it difficult to do. In the non-symmetry–intervention condition, however, we expected fairly good conformity to ΔP. There is a clash between non-symmetrical negation and intervention in this condition, because intervention is one of the key factors triggering analytic and comparative thought for identifying causation correctly (Lagnado, Waldmann, Hagmayer, & Sloman, 2007; Sloman & Lagnado, 2005).

Method

Participants

One hundred and twenty-six undergraduate students at the Ritsumeikan University participated for a partial course credit. Since six of them were excluded from analysis as described in *Results* section, 120 participants (60 pairs) were the subject for analysis (74 females, 46 males; mean age 21.2 years, SD 4.4).

Design

Participants were randomly assigned to one of four conditions in a 2 (non-symmetry/symmetry in negation) × 2 (observation/intervention) factorial design. The configuration of the cell frequency of stimuli that each participant observed was not fixed in advance, though occurrence provability of stimuli was controlled. Table 3 indicates the expected (i.e., theoretical) probabilities and the actually observed probabilities of stimuli in four sessions. Participants in the intervention conditions selected actions for the cause event at every trial, making the target cause present (*C*) or absent (¬*C*). The results of such an action, i.e., the effect present (*E*) or absent (¬*E*), was randomly determined each time according to the expected probabilities. The presentation order of these probabilities was randomised for participants.

Participants in the intervention conditions were allowed to select their action on cause events, and they observed the result that was determined according to the provability in every trial. They acted as many times as they liked, from minimum of 10 to maximum of 100. That is, the number of trials in a session was not fixed in advance in this experiment. The reason why the trial number was not fixed was to maximise people's freedom. In the real world, if intervention is possible, we usually intervene as many times as we like until we are happy to infer a conclusion. In order to guarantee equality of information that participants receive about the cell sequence (i.e., *a*, *b*, *c* or *d*), we adopted a *yoked* control design (see e.g., Moore & Gormezano, 1961; but see Church, 1989). Specifically, each participant in the observation conditions was paired (without their knowledge) with an arbitrarily defined participant in the intervention conditions, and given the exactly the same sequence of causal events as his or her counterpart.

Material and procedure

The task was to evaluate the causal effect on egg laying of a fictitious new feed product for hens. Participants were told to imagine that they were

Table 3. Expected (theoretical) and observed probabilities of stimuli and participants' judgements in Experiment 1.

	Expected (theoretical)				Observed							Judgement	
No.	$P(E\|C)$	$P(E\|\neg C)$	ΔP		N	$P(C)$	$P(E\|C)$	$P(E\|\neg C)$	$P(C\|E)$	ΔP	DFH	IV	OB
1	0.3	0.3	0	*M*	29.5	0.52	0.33	0.27	0.59	0.06	0.43	11.1	0.1
				SD	20.1	0.09	0.15	0.19	0.24	0.24	0.16	35.5	39.6
2	0.7	0.7	0	*M*	30.3	0.51	0.73	0.68	0.53	0.05	0.62	9.5	21.3
				SD	21.4	0.08	0.13	0.16	0.11	0.22	0.10	33.8	36.3
3	0.6	0.2	0.4	*M*	30.9	0.51	0.63	0.22	0.77	0.42	0.69	52.7	43.5
				SD	25.3	0.05	0.14	0.16	0.14	0.18	0.11	28.7	34.5
4	0.8	0.4	0.4	*M*	25.9	0.52	0.79	0.39	0.69	0.39	0.73	53.9	50.3
				SD	16.6	0.09	0.11	0.16	0.11	0.18	0.08	30.2	30.0

Note: *N* indicates the number of trials that participants performed.

poultry farmers, wondering whether they should introduce a new feed product for their hens in order to increase egg production. All participants took part in four sessions, each of which started with the learning phase, where they read the instructions (see Appendix), and observed a set of stimuli: information about hen-feeding and egg production sequentially presented (see Figure 1). They then went on to the evaluation phase, where they evaluated the causal influence of the target product on egg laying, on a scale from −100 (completely inhibited) to 100 (completely facilitated), by sliding a lever on the screen.

In the symmetrical negation conditions, the contrast was between two alternative hen feeds (i.e., *X* or *Y*): a new product developed by *X* company and a conventional product made by *Y* company that the farmer had been using so far. Figure 1(b) depicts an example of stimuli presentation (i.e., a *d*-cell case). In every trial, participants observed a hen that was given either *X* or *Y*, and then they were informed about the consequence: whether the hen laid an egg or not the next morning. Each hen was fed either *X* or *Y* only, so that feeding *Y* was equivalent to *not-X*. The following situations never occurred: a hen was given both foods (*X* and *Y*), a hen was given no food (*not-X* and *not-Y*), and a hen was given another type of food (*Z*). In the non-symmetrical negation conditions, the contrast was between having a dietary supplement at feeding time and not having one, i.e., *X* or *not-X*. Figure 1(c) depicts an example of such stimuli presentation (i.e., a *d*-cell case). On every trial, after observing whether a hen was given a supplement or not, participants were informed about whether or not it laid an egg the next morning.

In the intervention conditions, participants were able to choose their actions on the cause events. Figure 1(a) indicates a screen example for choosing in the symmetry–intervention condition. In this condition, participants chose either feed *X* or *Y* to give to a hen, and they then saw a stimulus of causal events, such as Figure 1(b), at every trial. In the non-symmetry–intervention condition, they selected whether or not they gave a dietary supplement to a hen at every trial in the same way. All participants in the intervention conditions were allowed to test samples (i.e., hen feeds or dietary supplement) as many times as they liked (the number of trials was actually limited between minimum of 10 and maximum of 100). The "evaluate" button, shown in Figure 1(a), appeared after they performed 10 trials, including at least both 1 present and 1 absent cases of the target cause events, and the participants were then in the evaluation phase. On the other hand, participants in the observation conditions were just exposed to the sequence of stimuli, such as (b) or (c), one by one, and never saw the choice screen, such as (a). The sequences of stimuli they observed were exactly the same as those that their counterparts observed in the intervention conditions. After they observed a sequence of stimuli, they automatically moved on to the evaluation phase.

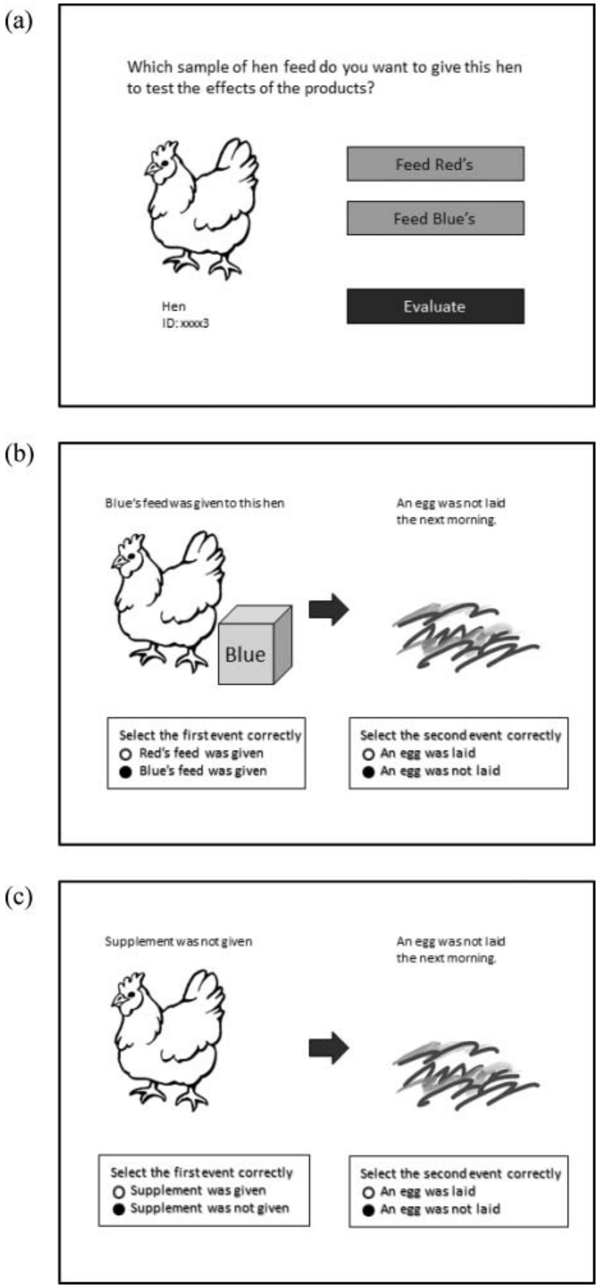

Figure 1. Examples of stimuli. Panel (a) illustrates an example of action-selecting screens that presented only in the intervention conditions (e.g., a case for the symmetry–intervention condition). The Evaluate button appeared only after participants performed 10 trials, including at least 1 present and 1 absent cases of the target cause. Panels (b) and (c) indicate examples of the screens that cell information was presented at trials. At each

Experiments were run in small groups of varying sizes (maximum of 9). Stimuli and full text instructions were presented on each monitor of the personal computers. After receiving brief oral instructions and a short exercise session, participants carried out the experiments at their own pace, changing the screens themselves. Each stimulus had four check boxes below the illustrated causal events (see Figure 1(b) and 1(c)). In order to make participants observe carefully which events happened, the next stimulus was presented only when participants checked the boxes correctly.

Result

The data obtained from five participants were eliminated from analysis, because two pairs included the participants who gave inappropriate answers (i.e., they answered 0 for all the stimulus sets), and one had a computer trouble.

For statistical analyses, we applied linear mixed-effects regression models (e.g., Baayen, Davidson, & Bates, 2008) using the lme4 (ver. 1.1.12, May 2016) and lmerTest (ver. 2.0.33, Dec 2016) on R (ver. 3.3.2, Oct 2016) to the data. Note that we cannot use ANOVA to analyse data from the current experiment, although the experimental design may seem to be a typical 2×2 factorial one. This is because the cell configuration of stimulus sets was different for each participant: the number of cases where the possible cause is present (sum of cells a and b) and it is absent (sum of cells c and d) depends on each participant's free will. For example, in the intervention conditions, a certain participant may feed 50 times (i.e., $a + b = 50$), but another participant may feed much fewer times.[3] Moreover, as the outcome of each participant's action is only determined probabilistically, realised probabilities based on actual frequencies are unlikely to match exactly the predefined ones shown in Table 3, e.g., it is possible that $a : b = 24 : 26$, $a/(a + b) = .48$, even though there is a theoretical probability of $P(E|C) = .5$. Analyses based on linear mixed-effects models, which is receiving more and more attention as a

trial, a cause event (i.e., a feeding chicken) appeared on screen, and then an arrow and an effect event (i.e., presence or absence of an egg) appeared in turn every one second. (b) and (c) depict the d-cell examples in the symmetry and non-symmetry conditions, respectively. The check boxes below the causal events were prepared in order to make participants observe carefully what events happened. The stimuli actually used were coloured and written in Japanese. The illustration of hen was adopted from free web contents (www.ActivityVillage.co.uk).

[3]The observed cell configurations of stimuli that each participant pair actually observed are available online.

powerful tool in various areas (e.g., Aarts, Verhage, Veenvliet, Dolan, & van der Sluis, 2014; Judd, Westfall, & Kenny, 2012), can be applied to this type of data, and we adopted this statistic method.

We compared the fitness of the following two linear mixed-effects regression models (in line with the notation of lme4) to participants' causal judgements in the case of DFH:

$$\text{M1 : Judgement} \sim \text{DFH} + (1|\text{Participant}) \tag{3}$$

$$\text{M2 : Judgement}{\sim}\text{DFH} + (\text{DFH} \,\|\, \text{Participant}) \tag{4}$$

M1 supposes Judgement is predicted by DFH, which is treated as a linear fixed effect, and by Participant, which is treated as a random effect (intercept). This means that every participant has a same effect (slope) of DFH, but has a different offset (intercept). According to M2, the slope and the intercept for the effect of DFH are determined independently for each level of Participant, assuming the random slopes and intercepts are independent. This means that every participant has a different effect (slope) of DFH and has a different offset (intercept). In the case of ΔP, everything is the same except that DFH is replaced by ΔP.[4]

Table 4 indicates results of the analyses. The Bayesian information criterion (BIC) was used to compare the predictability of the models; among four models, M1 and M2 for DFH and ΔP.[5] BIC is a measure of the relative goodness-of-fit of models that prescribes to select the model with the minimum value. Figure 2 shows the difference in the BIC (i.e., ΔBIC) between the best fit models for DFH and ΔP (either M1 or M2 for each). A positive value indicates that DFH outperformed ΔP, and a negative value indicates the opposite. DFH had a better fit with the data in the non-symmetry–observation condition, but ΔP explained the data better in the other three conditions.

Discussion

As predicted, participants' causal judgements were considerably affected by the factors of symmetrical negation and intervention. They were consistent

[4]There are two other possible mixed-effects models available: M0 and M3. M0 is the simplest model that has no fixed effect: Judgement $\sim$ 1 + (1 | Participant). This model assumes Judgment is predicted by neither DFH nor ΔP, and so we omit this model. M3 is the most complex model in which the slope and the intercept for the effect of DFH (or ΔP) are determined separately for each level of Participant as in the case of M2, while (unlike M2) allowing correlation between the intercept deviations and the effect of DFH (or ΔP) deviations within levels of Participant. Thus, M3 includes an additional parameter: the correlation between intercept deviations and DFH (or ΔP) deviations across levels of Participant. Although we prefer simpler models, we actually evaluated M3 against M1 and M2, and in most cases (seven out of all eight cases), M3 showed the worst fit to data according to BIC. Consequently, we decided to omit M3.

[5]Barr, Levy, Scheepers, and Tily (2013) proposed the maximal model should be the "gold standard" in model selection although it seems to be still controversial. In this study, M3 mentioned in the footnote 4 is the maximal model, and the results of Experiments 1 and 2 do not alter even using this standard.

Table 4. Summary of fixed effect predictors from the linear mixed-effects regression model for predicting causal strength in Experiment 1.

	Factor									
	Negation	Intervention	Model	BIC		Beta	Std. Error	df	*t*-Value	*Pr*(>|*t*|)
DFH										
	Non-symmetry	Observation	M1	1156.3	(Intercept)	−84.0	11.1	114.2	−7.5	.0000
					DFH	180.3	17.2	104.9	10.5	.0000
		Intervention	M1	1156.4	(Intercept)	−47.0	11.2	116.0	−4.2	.0001
					DFH	123.9	17.6	116.0	7.0	.0000
	Symmetry	Observation	M1	1184.0	(Intercept)	−58.1	9.6	118.7	−6.1	.0000
					DFH	142.9	14.6	105.0	9.8	.0000
		Intervention	M1	1222.2	(Intercept)	−56.3	11.5	124.0	−4.9	.0000
					DFH	145.9	17.9	124.0	8.1	.0000
Δ*P*										
	Non-symmetry	Observation	M1	1189.9	(Intercept)	7.1	4.5	58.2	1.6	.1189
					Δ*P*	91.4	12.7	110.2	7.2	.0000
		Intervention	M2	1116.3	(Intercept)	5.8	3.0	102.3	1.9	.0599
					Δ*P*	108.1	11.3	39.2	9.6	.0000
	Symmetry	Observation	M1	1147.4	(Intercept)	7.6	3.2	55.5	2.4	.0222
					Δ*P*	95.2	7.2	108.0	13.2	.0000
		Intervention	M2	1197.6	(Intercept)	10.8	3.2	43.4	3.4	.0014
					Δ*P*	101.2	11.1	44.7	9.1	.0000

Note: *Model* indicates which model of the two linier mixed-effects regression models (i.e., M1 and M2, see text in detail) was selected.

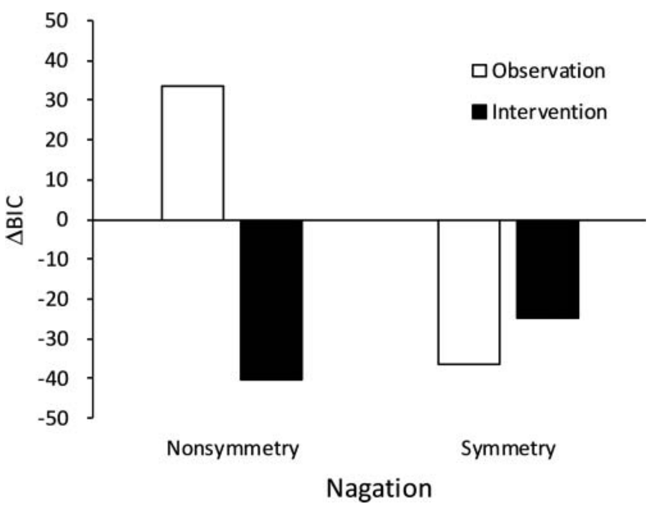

Figure 2. Relative fit of DFH (A-frame) against Δ*P* (B-frame) based on BIC for predicting causality judgements as a function of two factors (symmetry in negation and intervention) in Experiment 1. Positive value indicates better fitness for DFH, and negative one indicates better fitness for Δ*P*.

with DFH in the non-symmetry–observation condition, but with Δ*P* in the symmetry–intervention condition and the other two conditions. That is, participants disregard the *d*-cell information in the non-symmetry–observation condition, as many previous studies have reported. They did, however, take account of the *d*-cell in the other three conditions in which these factors were introduced. The results demonstrated that normative ΔP, which have been criticised repeatedly for its discrepancy with experimental data, was more descriptive in the conditions where symmetry and/or intervention factors existed. The results suggest that participants use two frames to make causal judgements, depending on condition. Using the B-frame, participants tended to pay attention to the *d*-cell as well as the other cell information.

The results supply evidence that these two factors, symmetrical negation and intervention, encouraged participants to switch from the default A-frame to the B-frame. In the symmetry conditions, the occurrence of the alternative causal event (i.e., Y) was, in effect, the negation of the target cause (i.e., $\neg X$). This alternative to the target event was likely to stand out and not be a mere "ground" for a "figure" in this situation, and accordingly, the frame shift from A to B became more probable. Switching to the B-frame, the participants focused more equally on the positive and negative features. Therefore, ΔP outperformed DFH in the symmetry conditions, because participants tended to consider the information from all the cells.

The results also demonstrated the remarkable effectiveness of intervention for shifting the frame. When participants were able to make an intervention, their causal judgements were consistent with ΔP, whether negation was symmetry or non-symmetry. Non-symmetry in negation is compatible with "figure–ground" relationship, where the non-occurrence of events is prone to be disregarded as "ground" information. That it, ΔP should not have described causal judgements in the non-symmetry conditions. However, intervention overrode this effect and helped participants to switch to the B-frame, where both the occurrence and non-occurrence of events were focused on, with the result that *d*-cell information was taken into account. The results imply that intervention is a powerful factor for changing what participants find salient and relevant in causal judgement, as Anderson and Sheu (1995) argued.

In Experiment 2, we investigated the influence of symmetrical negation and commitment. Since the influence of intervention was clearly shown to be present in Experiment 1, we focused only on observation in Experiment 2.

Experiment 2

In Experiment 2, the influence of symmetrical negation (non-symmetry vs. symmetry) and commitment (low-commitment vs. high-commitment) were examined. Trial numbers were fixed in advance this time. Therefore, the configuration of stimuli sets was common to all participants in this experiment, unlike in Experiment 1. We predicted that participants' causal judgements would be consistent with DFH in the non-symmetry–low-commitment condition and with ΔP in the symmetry–high-commitment condition, because

Table 5. Frequency distribution of stimuli in Experiment 2.

	Cell configuration					Model		Bet
No.	*a*	*b*	*c*	*d*	*N*	Δ*P*	DFH	Outcome
1	1	4	4	4	13	−0.30	0.20	lose
2	4	4	4	1	13	−0.30	0.50	lose
3	1	3	4	12	20	0.00	0.22	lose
4	5	5	5	5	20	0.00	0.50	draw
5	12	4	3	1	20	0.00	0.77	win
6	2	0	6	6	14	0.50	0.50	win
7	9	0	5	5	19	0.50	0.80	win

symmetrical negation and commitment are factors that encourage the switch from the A-frame to the B-frame. For the other two conditions, interactions between the two factors were expected. These conditions include either one of the factors, symmetrical negation or high-commitment. However, some participants may still find it difficult to shift to the B-frame, since it demands high cognitive resources.

Method

Participants

A total of 102 undergraduate students (female 43, male 59) of the Ritsumeikan University participated in the experiment for a partial course credit (mean age 19.9 years, SD 0.9).

Design

Participants were randomly assigned to one of four conditions in a 2 (non-symmetry/symmetry in negation) × 2 (low-/high-commitment) factorial design. Table 5 indicates the configuration of the cell frequency of stimuli that each participant was shown in seven sessions. This configuration was designed to provide a significant difference between Δ*P* and DFH.[6] The presentation order was counterbalanced among participants.

Material and procedure

The task was again to evaluate the causal influence of fictitious feed products on egg laying. Participants were to imagine, as in Experiment 1, being poultry

[6]We selected these four stimuli because there should be multiple levels for Δ*P*, and there also should be a variation in other indices that have the same Δ*P* value. There was no stimulus that had high Δ*P*, since it was difficult for high Δ*P* to make variations in combination of $P(E|C)$ and $P(E|\neg C)$. For example, to set Δ*P* 0.8, only a combination of 0.9 and 0.1 is available in a grid scale of 0.1, if an extreme probability of 1 or 0 is avoided. Moreover, such stimuli do not provide a significant difference between DFH and Δ*P*. Using a computer program, we exhaustively searched for a combination of stimuli that has a lesser internal correlation between Δ*P* and DFH. Under a constraint that the sample size is equal to or fewer than 20, a set of stimuli that have 3 (Δ*P*: low/middle/high) × 3 (DFH: low/middle/high) levels was searched, but there was no such stimulus that had a high–low or low–high combination in Δ*P* and DFH. As a consequence, the levels of Δ*P* were set at approximately −.3, .0 and .5, and the levels of DFH were set at approximately .2, .5 and .8.

farmers who were thinking about introducing a new feeding product to increase egg laying in their hens, but this time there was no intervention, only observation. The participants only observed the causal events as test results at a neighbouring poultry farm. All participants performed in seven sessions, each of which consisted of the learning phase, reading an instruction and observing a set of stimuli causal events sequentially, and the evaluation phase, rating the causal influence of the target product on egg laying on a scale from −100 to 100. Preventive causality was taken into account in this experiment. It was emphasised, in the instructions, that the target products might have a negative effect on the occurrence of the target effect. In the evaluation phase, participants first selected one of the two options about effectiveness – facilitative or inhibitive – on the screen, and then they indicated the degree of the effect, on a scale from 0 to 100, by sliding a lever presented below the options. To indicate no effectiveness, they were to select facilitative option with a degree of 0.

In Experiment 2, the symmetrical negation of an effect was also manipulated, in addition to that of a cause. The negation symmetry of causal events was manipulated in the same way as Experiment 1, where the contrast was between giving a new supplement (the target cause) and not giving it in feeding in the non-symmetry conditions, or between giving a new feed *X* (the target cause) and giving an accustomed feed *Y* (the alternative cause). A parallel contrast was introduced for the possible effects in Experiment 2: a contrast between laying an egg (the target effect) and not laying an egg in the non-symmetry conditions, or between laying a brown egg with high value in the market (the target effect) and laying a white egg with low value (the alternative effect) in the symmetrical conditions. To summarise, in the symmetrical conditions, both the possible cause and effect had symmetrical negations. There were two kinds of hen feed, *X* or *Y*, for causes, with one a new product, the target cause, and the other the one already in use. There was also the possible production of brown eggs (i.e., the target effect) or white eggs (i.e., the alternative effect). Participants were told that brown eggs were popular with consumers, making it advantageous to produce brown eggs. They saw which feed a hen was given, and then saw on the same screen whether a brown or white egg was laid the next morning. The non-symmetry conditions were same as Experiment 1. The participants saw whether or not a hen was given the supplement during feeding, and then saw whether or not an egg was laid the next morning.

In order to manipulate the degree of commitment, betting on one's own judgements, followed by feedback, was introduced. In the high-commitment conditions, before all of the sessions began, participants were given points to the value of 7000 yen (approx. $70) and asked to increase this value as much as possible by betting at the evaluation phases. In every session, after they evaluated the causal influence of the target feeding products, they were asked to bet. The bet was about whether the neighbouring farm should adopt the new feeding product (i.e., the target cause) to increase profit. There were three options: (a) betting 1000 yen (approx. $10) in favour of the new feeding product, (b) betting 1000 yen against it, and (c) not betting on this occasion. After participants made their choice, they were given feedback on their betting (win, lose or draw) and their resulting balance. When either ΔP or DFH was high (see Table 5), the profit increased with option (a) winning, or the profit decreased with option (b) losing. When either ΔP or DFH was low, the profit decreased with option (a) losing, and the profit increased with option (b) winning. When ΔP and DFH were in the middle range, there was no change in profit, and the bet was a draw. Winning participants got 1000 yen added in points, while losing participants had 1000 yen deducted in points. Participants neither won nor lost points when they selected (c) and the bet was a draw. In the low-commitment conditions, participants just observed the events and evaluated the causal influence without betting.

Experiments were run in small groups of varying sizes (maximum of 24). Stimuli and full text instructions were presented on each monitor of the personal computers. After receiving brief oral instructions and a short exercise session, participants worked at their own pace, changing the screens themselves.

Results and discussion

For a comparison with the results of Experiment 1, the same procedures for analysis were adopted: the same linear mixed-effects regression models were applied to the data. As shown in Table 6, DFH showed good fit with data in both non-symmetry conditions. ΔP predicted the result better than DFH in the symmetry conditions, but the difference of BIC values between DFH and ΔP was small in the symmetry–low-commitment condition.

Figure 3 shows a comparison of BIC between ΔP and DFH. The meaning of the sign ΔBIC is the same as in Experiment 1. The influence of symmetrical negation was similar to that in Experiment 1. Again, ΔP corresponded more with causal judgements in the symmetrical conditions than with judgements in the non-symmetry conditions, and DFH matched the non-symmetry conditions rather than the symmetrical conditions.

There were also effects of commitment on causal judgements. ΔP showed better fit in the high-commitment conditions, though the effects of commitment were not so strong as those of intervention in Experiment 1. One possibility is that the degree of commitment controlled by betting did not exceed a certain point for some participants. We used participants' bets of fictitious money on their judgements to control the commitment factor in this experiment. It is plausible that the gain and loss of fictitious money might not have increased commitment enough for some participants to make the frame shift.

Table 6. Summary of fixed effect predictors from the linear mixed-effects regression model for predicting causal strength in Experiment 2.

	Factor									
	Negation	Commitment	Model	BIC		Beta	Std. Error	*df*	*t*-Value	*Pr*(>|*t*|)
DFH										
	Non-symmetry	Low commitment	M1	1908.8	(Intercept)	−46.8	6.5	183.0	−7.2	.0000
					DFH	121.1	11.6	162.0	10.4	.0000
		High commitment	M1	1747.9	(Intercept)	−57.7	8.0	168.0	−7.2	.0000
					DFH	155.7	14.6	168.0	10.6	.0000
	Symmetry	Low commitment	M1	1781.1	(Intercept)	−41.8	7.0	173.2	−6.0	.0000
					DFH	117.7	12.7	150.0	9.3	.0000
		High commitment	M1	1878.7	(Intercept)	−33.3	7.4	181.0	−4.5	.0000
					DFH	92.6	13.5	156.0	6.9	.0000
ΔP										
	Non-symmetry	Low commitment	M1	1955.9	(Intercept)	10.3	3.0	28.8	3.4	.0019
					ΔP	60.6	9.6	162.0	6.3	.0000
		High commitment	M1	1778.4	(Intercept)	14.8	3.6	168.0	4.2	.0000
					ΔP	93.1	11.4	168.0	8.2	.0000
	Symmetry	Low commitment	M1	1770.0	(Intercept)	12.0	3.0	26.4	4.0	.0005
					ΔP	88.9	8.7	150.0	10.2	.0000
		High commitment	M2	1851.5	(Intercept)	8.4	3.1	26.5	2.7	.0115
					ΔP	80.2	11.9	26.5	6.7	.0000

Note: *Model* indicates which model of the two linier mixed-effects regression models (i.e., M1 and M2, see text in detail) was selected.

Analytic thought demands high cognitive resource (e.g., De Neys, 2006) especially when negation is non-symmetrical, and when that is so, the prospect of mere fictitious money might not have been enough to encourage some participants to switch to the B-frame.

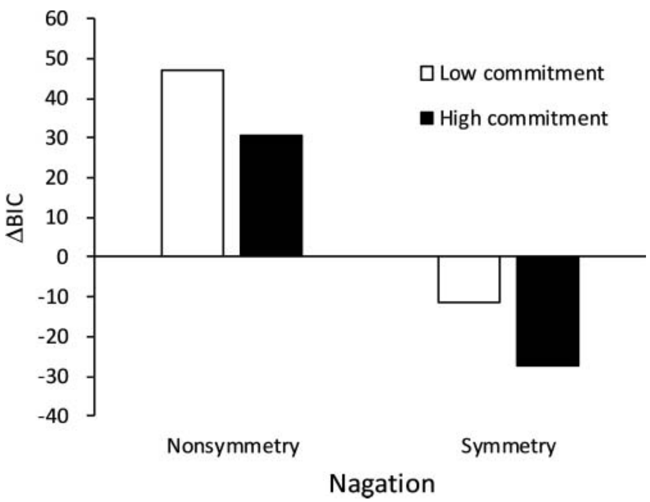

Figure 3. Relative fit of DFH (A-frame) against ΔP (B-frame) based on BIC for predicting causality judgements as a function of two factors (s symmetry in negation and commitment) in Experiment 2. Positive value indicates better fitness for DFH, and negative one indicates better fitness for ΔP.

General discussion

Our predictions about DFH and ΔP, in the specified conditions, were consistent with the participants' causal inductions in the findings of Experiments 1 and 2. DFH fit the conditions where figure–ground framing was likely to arise, that is, when negation was non-symmetrical, intervention was not allowed, and commitment was not high enough. In contrast, ΔP was more descriptive of people's causal judgements where the comparative view was likely to arise, that is, when negation was symmetrical, intervention was allowed, and commitment was high enough. These findings support the hypotheses that these factors – symmetry in negation, intervention and commitment – encourage frame shifting from the default A-frame to the B-frame, and that people use two different frames depending on conditions. In the B-frame, represented by ΔP, negative features as well as positive features are taken into consideration: both occurrence and non-occurrence are focused on for comparison. When the B-frame was activated by its relevant factors, participants considered the *d*-cell in addition to the other cells, with the result that they conformed to ΔP. On the other hand, in the A-frame, represented by DFH, positive features in events were attended to: occurrence was focused on and non-occurrence was ignored. Thus when the A-frame was activated by its relevant factors, participants were prone to disregard *d*-cell information.

The present findings show that there is a psychological asymmetry between occurrence and non-occurrence. When negation is symmetrical, that is $\neg X$ can be labelled as Y, the two alternatives, X and Y, seem to be on psychologically even ground, and switching to the B-frame, from the default A-frame, is easier. Then *d*-cell information, which is often ignored in common causal contexts, is taken into account along with the other cells, and ΔP becomes more descriptive. In Experiment 1, it was found that ΔP showed high correlations with causal judgements in the symmetry as opposed to the non-symmetry conditions. DFH showed the reverse pattern, indicating better fit with causal judgements in the non-symmetry conditions, when the possible cause C had a relatively small set size. The results in Experiment 2 were also consistent with prediction when the distinction between symmetrical and non-symmetrical negations was extended to the effect events. ΔP showed a better fit in the symmetry conditions than in the non-symmetry conditions, but DFH was better in the non-symmetry conditions than in the symmetry conditions. Results similar to this psychological asymmetry between affirmation and negation (i.e., X and $\neg X$) in causal reasoning have been found in other areas. For example, Oaksford and Stenning (1992) showed that replacing a negated constituent with super-ordinate category (e.g., vowel vs. consonant instead of A vs. not-A) reduced matching bias in the selection task (see also Hattori, 2002; Oaksford, 2002). See Hattori et al. (2016) for further details.

These results support the psychological validity of the rarity assumption in people's judgements about causation (Hattori, 2002; McKenzie & Mikkelsen, 2000, 2007; Oaksford & Chater, 1994, 2003). In standard causal contexts, instances of the possible cause *C*, such as eating sushi at a particular restaurant, would be "rare" relative to $\neg C$, not eating sushi. The occurrence of the effect *E*, such as feeling unwell, would be "rare" relative to $\neg E$, not feeling unwell. It is often efficient to have a heuristic that is based on DFH, ignoring the *d*-cell, rather than a process that conforms precisely to ΔP, which takes account of the *d*-cell. However, when logical negation can be replaced by symmetrical negation, as when no sushi eating is replaced by tempura eating, it is efficient and more accurate to conform to ΔP. In our terms, the switch from the DFH heuristic to the ΔP relation is a change from the A-frame to the B-frame. Other approaches, such as the focal set analysis of Cheng (1997), might provide a partial explanation of some of our results, but the effects of symmetry in negation would be hard to explain.

The findings in Experiment 1 also imply that intervention is crucial for changing the frames. Intervention serves not only to identify the correct causal structure among events (Lagnado et al., 2007; Sloman & Lagnado, 2005), but also to change what is salient and relevant in causal judgement (Anderson & Sheu, 1995). With the B-frame, there is equal focus on positive and negative events, so that the figure–ground effect is less likely to occur. Consequently, in the intervention conditions, ΔP, where the *d*-cell information was made use of, had a good fit with participants' causal judgements. Even if negation was non-symmetrical, where the figure–ground framing was apt to arise, participants who were allowed interventions made judgements consistent with ΔP. In the non-symmetry–observation condition, DFH, where the *d*-cell information was disregarded, had better fit with the data.

Commitment also had an effect on fame shifting. Although it was not as strong as intervention, commitment, with its associated benefits and costs, is likely to be relevant to frame shifting. The findings in Experiment 2 suggest that commitment influences the effectiveness of negation symmetry. Probably it was difficult for some participants to use the B-frame without commitment even if there was symmetrical negation, since the B-frame demands high cognitive load. It suggests that the factors for frame shifting probably work additively: the more relevant factors there are, the more likely it becomes that there will be a switch to the B-frame.

The results of the current experiments reconcile normative and descriptive views of causal reasoning. Lu, Yuille, Liljeholm, Cheng, and Holyoak (2008, p. 976) criticised DFH as being non-normative. The grounds of their criticism are (1) the phi coefficient on which DFH is based does not make reference to underlying causal relationships, (2) DFH ignores *d*-cell information, and (3) DFH does not properly handle preventive causes. What is overlooked in this criticism is that DFH is a model of only one of the two stages of causal induction. There is the heuristic stage for narrowing the target, and the analytic stage for more controlled process (Hattori & Oaksford, 2007, p. 767), corresponding to the A- and B-frames, respectively. DFH is only a model of the A-frame. It does not need to detect "underlying causal relationships", but rather to reveal correlations, and it saves cognitive resources by disregarding the *d*-cell.

The reason why DFH does not cover preventive causes lies in the psychological asymmetry between affirmation and negation. It is often important to detect an affirmative cause like catching a cold, but it is less urgent to discover a negative cause, like being immune to the cold virus. There must be a close relation to the way we communicate: our verbal communication must have been constructed in such a manner that important information is conveyed in the affirmative rather than in negative. Return to our example of an outbreak of food poisoning. It would be exceedingly inefficient to compare every predictive probability of food poisoning using the B-frame, listing Kamikaze roll sushi at sushi bar 1, tuna sushi at sushi bar 2, some kind of mushrooms at restaurant 3, and so on. On the other hand, the use of DFH for quickly and efficiently inferring a tentative causal hypothesis can be beneficial. It can help us to decide what to do immediately, such as cancelling the reservation at a restaurant under suspicion. It is highly unlikely that there is a single "golden" (and normative) model of all aspects of causal reasoning.

Another criticism of DFH by Lu et al. (2008) about descriptive validity has methodological implications. In all their experiments, covariation information was presented in the same way, using two pictures, indicating frequencies of the effect present given the cause present and given the cause absent. The two pictures always contained the same number of instances. This presentation can motivate participants to compare two predictive probabilities, $P(E|C)$ and $P(E|\neg C)$, which is a differentiation mode comparable with the B-frame (see Table 2). Since DFH is a model of the A-frame, however, it is plausible that this type of presentation decreases its model fit. It would be worth investigating the effect of different stimulus presentations on frame switching.

Finally, our findings on the A- and B-frames and causal induction provide broad support for general dual process theories of human thought (e.g., Evans, 2003; Evans & Over, 1996; Sloman, 1996; Stanovich, 1999). In these theories, there are two kinds of mental processes, one heuristic and one analytic. The former is said to be of type 1 or in System 1, and the latter of type 2 or in System 2. The A-frame corresponds to heuristic processing, using DFH, and the B-frame to a more analytical thought, conforming to ΔP. Consistent with points made by Evans and Stanovich (2013), we do not argue that the heuristic A-frame process is necessarily "biased" and inferior to an analytic B-frame process, nor that the latter is necessarily more "rational" than the former. The A-frame can be more efficient than the B-frame in certain circumstances, and the B-frame can waste time and energy on an unnecessarily high level of accuracy in some contexts. Of course, the increased accuracy that comes with the B-frame can sometimes be worth its extra cost in cognitive resources, but it will not be worthwhile at other times, when the A-frame can allow us to make speedy and efficient judgements, as a default mode of thinking. However, we leave for future research the question of exactly how our frames are related to the distinction between type 1 and type 2 processes, or between System 1 and System 2, in dual process theories.

Acknowledgments

We thank Yutaka Nishida and Henrik Singmann for their valuable comments on data analyses, Masahiko Tamura for his coding computer programs for experiments, Ryo Orita for his assistance in conducting the experiments, and the members of CHORUS project for exciting discussions on our earlier study.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

JSPS–ANR CHORUS Program [grant number J12100148]; JSPS KAKENHI awarded to MH [grant number 22500247], [grant number 15H02717].

ORCID

Ikuko Hattori http://orcid.org/0000-0003-0196-9764
Masasi Hattori http://orcid.org/0000-0002-6974-6906
David E. Over http://orcid.org/0000-0002-0353-9381
Tatsuji Takahashi http://orcid.org/0000-0002-8891-5664
Jean Baratgin http://orcid.org/0000-0001-9566-486X

References

Aarts, E., Verhage, M., Veenvliet, J. V., Dolan, C. V., & van der Sluis, S. (2014). A solution to dependency: Using multilevel analysis to accommodate nested data. *Nature neuroscience, 17*, 491–496. doi:10.1038/nn.3648

Anderson, J. R., & Sheu, C.-F. (1995). Causal inferences as perceptual judgments. *Memory & Cognition, 23*, 510–524. doi:10.3758/bf03197251

Baayen, R. H., Davidson, D. J., & Bates, D. M. (2008). Mixed-effects modeling with crossed random effects for subjects and items. *Journal of Memory and Language, 59*, 390–412. doi:10.1016/j.jml.2007.12.005

Barr, D. J., Levy, R., Scheepers, C., & Tily, H. J. (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language, 68*, 255–278.

Cheng, P. W. (1997). From covariation to causation: A causal power theory. *Psychological Review, 104*, 367–405. doi:10.1037/0033–295X.104.2.367

Church, R. M. (1989). The yoked control design. In T. Archer & L. G. Nilsson (Eds.), *Aversion, avoidance, and anxiety: Perspectives on aversively motivated behavior* (pp. 403–415). Hillsdale, NJ: Lawrence Erlbaum Associates.

De Neys, W. (2006). Dual processing in reasoning: Two systems but one reasoner. *Psychological Science, 17*, 428–433. doi:10.1111/j.1467-9280.2006.01723.x

Evans, J. St. B. T. (2003). In two minds: Dual-process accounts of reasoning. *Trends in Cognitive Sciences, 7*, 454–459. doi:10.1016/j.tics.2003.08.012

Evans, J. St. B. T., & Over, D. E. (1996). *Rationality and reasoning*. Hove: Psychology Press.

Evans, J. St. B. T., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*, 223–241. doi:10.1177/1745691612460685

Fitelson, B., & Hitchcock, C. (2011). Probabilistic measures of causal strength. In P. M. Illari, F. Russo, & J. Williamson (Eds.), *Causality in the sciences* (pp. 600–627). Oxford: Oxford University Press.

Griffiths, T. L., & Tenenbaum, J. B. (2005). Structure and strength in causal induction. *Cognitive Psychology, 51*, 334–384. doi:10.1016/j.cogpsych.2005.05.004

Hattori, M. (2002). A quantitative model of optimal data selection in Wason's selection task. *The Quarterly Journal of Experimental Psychology A, 55*, 1241–1272. doi:10.1080/02724980244000053

Hattori, M. (2014). Figure and ground in thinking: The affirmation-negation asymmetry as a consequence of framing. *Ritsumeikan Journal of Human Sciences, 636*, 131–147.

Hattori, M., & Oaksford, M. (2007). Adaptive non-interventional heuristics for covariation detection in causal induction: Model comparison and rational analysis. *Cognitive Science, 31*, 765–814. doi:10.1080/03640210701530755

Hattori, M., Over, D. E., Hattori, I., Takahashi, T., & Baratgin, J. (2016). Dual frames in causal reasoning and other types of thinking. In N. Galbraith, E. Lucas, & D. E. Over (Eds.), *The thinking mind: A festschrift for Ken Manktelow*. Hove: Psychology Press.

Jenkins, H. M., & Sainsbury, R. S. (1969). The development of stimulus control through differential reinforcement. In N. J. Mackintosh & W. K. Honig (Eds.), *Fundamental issues in associative learning*. Halifax: Dalhousie University Press.

Jenkins, H. M., & Sainsbury, R. S. (1970). Discrimination learning with the distinctive feature on positive or negative trials. In D. I. Mostofsky (Ed.), *Attention: Contemporary theory and analysis*. New York, NY: Appleton-Century-Crofts.

Jenkins, H. M., & Ward, W. C. (1965). Judgment of contingency between responses and outcomes. *Psychological Monographs: General and Applied, 79*, 1–17. doi:10.1037/h0093874

Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*, 54. doi:10.1037/a0028347

Kao, S.-F., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 19*, 1363–1386. doi:10.1037/0278-7393.19.6.1363

Lagnado, D. A., Waldmann, M. R., Hagmayer, Y., & Sloman, S. A. (2007). Beyond covariation: Cues to causal structure. In A. Gopnik & L. Schulz (Eds.), *Causal leaning: Psychology, philosophy, and computation* (pp. 154–172). New York, NY: Oxford University Press.

Lu, H., Yuille, A. L., Liljeholm, M., Cheng, P. W., & Holyoak, K. J. (2008). Bayesian generic priors for causal learning. *Psychological Review, 115*, 955–984. doi:10.1037/a0013256

McKenzie, C. R. M., & Mikkelsen, L. A. (2000). The psychological side of Hempel's paradox of confirmation. *Psychonomic Bulletin and Review, 7*, 360–366. doi:10.3758/BF03212994

McKenzie, C. R. M., & Mikkelsen, L. A. (2007). A Bayesian view of covariation assessment. *Cognitive Psychology, 54*, 33–61. doi:10.1016/j.cogpsych.2006.04.004

Moore, J. W., & Gormezano, I. (1961). Yoked comparisons of instrumental and classical eyelid conditioning. *Journal of Experimental Psychology, 62*, 552–559. doi:10.1037/h0044551

Oaksford, M. (2002). Contrast classes and matching bias as explanations of the effects of negation on conditional reasoning. *Thinking & Reasoning, 8*, 135–151. doi:10.1080/13546780143000170

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review, 101*, 608–631. doi:10.1037/0033-295X.101.4.608

Oaksford, M., & Chater, N. (2003). Optimal data selection: Revision, review, and reevaluation. *Psychonomic Bulletin and Review, 10*, 289–318. doi:10.3758/BF03196492

Oaksford, M., & Stenning, K. (1992). Reasoning with conditionals containing negated constituents. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*, 835–854. doi:10.1037/0278-7393.18.4.835

Pearl, J. (2000). *Causality: Models, reasoning, and inference*. Cambridge: Cambridge University Press.

Rubin, E. (1915/1958). Figure and ground. In D. C. Beardslee & M. Wertheimer (Eds.), *Readings in perception* (pp. 194–203). Princeton, NJ: D. Van Nostrand.

Rubin, E. (1921). *Visuell wahrgenommene figuren: Studien in psychologischer analyse* [Visually perceived figures: Studies in psychological analysis]. Kobenhavn: Gyldendalske Boghandel.

Schustack, M. W., & Sternberg, R. J. (1981). Evaluation of evidence in causal inference. *Journal of Experimental Psychology: General, 110*, 101–120. doi:10.1037/0096-3445.110.1.101

Sloman, S. A. (1996). The empirical case for two systems of reasoning. *Psychological Bulletin, 119*, 3–22. doi:10.1037/0033-2909.119.1.3

Sloman, S. A., & Lagnado, D. A. (2005). Do we "do"? *Cognitive Science, 29*, 5–39. doi:10.1207/s15516709cog2901—2

Spirtes, P., Glymour, C., & Scheines, R. (1993). *Causaton, prediction, and search*. New York, NY: Springer.

Stanovich, K. E. (1999). *Who is rational? Studies of individual differences in reasoning*. Mahwah, NJ: Elrbaum.

Waldmann, M. R. (1996). Knowledge-based causal induction. In D. R. Shanks, K. J. Holyoak, & D. L. Medin (Eds.), *The psychology of leaning and motivation, Vol. 34: Causal leaning* (pp. 47–88). San Diego, CA: Academic Press.

Appendix

Instructions in Experiment 1

Intervention conditions

Imagine that you are a poultry farmer who breeds hens and sells their eggs for living. Your hens are an improved variety called *Black Sander*, whose eggs are very tasty and rich in nutrients. Their eggs can be sold at a higher price than those of conventional hens, but it is uncertain whether a Black Sander hen lays an egg every morning. The current egg production rate of your hens is not very good.

(Symmetry condition) There are only two companies that provide hen feed in your country: *the Red* and *the Blue Corporations*. You have used the feed produced by the Blue Co. so far. You have heard that the hen feed produced by Red Co. might increase egg production. The change of hen feed from Blue's to the alternative Red's, however, costs you extra money. Therefore, you have decided to investigate the effectiveness of the new hen feeds to increase the egg production rate. Both companies, *Red* and *Blue*, have given you samples of their new products (i.e., hen feed) for a sales promotion. You have 100 samples of each product. You can give the samples to your breeding 100 hens to investigate to what degree the products affect egg production. You can use the samples as many times as you like. You can test both kinds of samples (hen feed) in the same numbers, or you can be biased to either one of them. Please note that you must feed a hen one of the two samples, but you cannot feed both samples to the same hen, due to possible side effects caused by compounding the ingredient in both products.

You do not need to test the samples on all of your 100 hens. When you think you have obtained enough data for an assessment, please stop examining and evaluate to what degree the new hen feed provided by *Red Co.* influences egg production. Please note that you have to test the samples on at least 10 hens before you evaluate the influence of the change. Moreover, you have to try at least one sample from each company.

(Non-symmetry condition) You have given only hen feed to your hens so far, but you have heard that giving vitamin supplements may increase the egg production rate of hens. It costs you extra money to give vitamin supplements to hen. Therefore, you have decided to investigate the effectiveness of the supplements for increasing the egg production rate. Currently, there is only one kind of vitamin supplement for hens provided by *the Yellow Co.* in your country. It has given you 100 samples of its new products (i.e., supplement) for a sales promotion. You can give the samples to your breeding 100 hens to investigate to what degree the products affect egg production. You can use the samples as many times as you like. You can test all samples (vitamin supplement) on your 100 hens, or you can give them to some of the hens and not others.

You do not need to test the samples on all of your 100 hens. When you think you have obtained enough data for an assessment, please stop examining and evaluate what degree the addition of the supplement influences egg production. Please note that you have to examine at least 10 hens before you evaluate the influence of the change. Moreover, you have to try at least one sample.

Observation conditions

Imagine that you are a poultry farmer who breeds hens and sells their eggs for living. Your hens are an improved variety called *Black Sander*, whose eggs are very tasty and rich in nutrients. Their eggs can be sold at a higher price than those of conventional hens, but it is uncertain whether a Black Sander hen lays an egg every morning. The current egg production rate of your hens is not very good.

(Symmetry condition) There are only two companies that provide hen feed in this country: *the Red* and *the Blue Companies*. You have used the feed produced by the *Blue Co.* so far. You have heard that the hen feed produced by the *Red Co.* might increase egg production. The change of hen feed from the Blue's to the alternative Red's, however, costs you extra money. Therefore, you have decided to investigate the effectiveness of the new hen feed to increase the egg production rate. Both companies, *the Red* and *the Blue*, have given 100 samples of their new products (i.e., hen feed) to your neighbouring poultry farmer for a sales promotion. Your neighbour also breeds 100 *Black Sander* hens and is trying to investigate the effects of hen feed on egg production. Please observe the tests of samples at your neighbouring farm carefully and then evaluate to what degree a change to the new hen feed provided by *the Red Co.* influences egg production.

(Non-symmetry condition) You have given only hen feed to your hens so far, but you heard that giving vitamin supplements might increase the egg production rate of hens. It costs you extra money to give vitamin supplements to hen. Therefore, you have decided to investigate the effectiveness of the supplements to increase the egg production rate. Currently, there is only one kind of vitamin supplement for hens provided by *the Yellow Co.* in your country. It has given 100 samples of their new products (i.e., supplement) to your neighbouring poultry farmer for a sales promotion. Your neighbour also breeds 100 *Black Sander* hens and is trying to investigate the effects of supplement on egg production. Please observe the tests of samples at your neighbouring farm carefully and then evaluate to what degree the addition of the supplement influences egg production.