# The Negative Contingency Illusion: A Cognitive Bias Leading to Misjudgement of Protection

**Richard M. Piech** ( richard.piech@uis.no )
University of Stavanger
**Matthew J. C. Crump**
Brooklyn College of the City University of New York
**David H. Zald**
Vanderbilt University Medical School

**Research Article**

**Keywords:** Contingency, Cognitive Bias, Misjudgement of Protection, Humans, world, judgements, actual relationship, outcome

**Posted Date:** July 9th, 2021

**DOI:** https://doi.org/10.21203/rs.3.rs-671507/v1

# Abstract

Humans possess a highly adaptive ability to draw inferences about the world by recognizing meaningful links between stimuli and events: making contingency judgements. We describe a systematic bias in contingency judgements that we label the *negative contingency illusion* in which individuals falsely judge a cue to be protective against an outcome. We demonstrate that the illusion arises when outcome probability is low and occurs when there is no actual relationship between cue and outcome and even when there is a modest positive relationship between cue and outcome. Such misjudgements may lead individuals to superstitious beliefs and could have major public health implications if they lead to the belief in and promotion of treatments that are ineffective or deleterious to the prevention and treatment of illness.


# Main Text

The ability to recognize meaningful associative links between stimuli or events is fundamental to drawing inferences about the world, predicting outcomes, and selecting actions based on cues in the environment (1). This recognition of associative links, or contingencies, and their strength, is highly adaptive. Humans have a strong propensity to make contingency judgements, which occur frequently and often automatically (2). However, errors in such judgements can have profound consequences. For instance, during the COVID-19 pandemic, errors in judging the strength of contingencies for getting ill or preventing illness have had broad life and death implications, impacting public health policies, medical treatment, economic policy, and countless individual decisions about how strictly to follow public health warnings.

The discrete-trials procedure is the most common laboratory method for studying contingency judgements (3-7). In this paradigm, a series of cue-outcome pairings (cue-outcome; cue-no outcome; no cue-outcome; no cue-no outcome) are presented, followed by a rating of the contingent relationship that was observed between the cue and outcome. The frequencies of the four event pairs can be manipulated to allow modulation of contingency strengths between the cue and the outcome, ranging from perfectly negative, totally unrelated, to perfectly positive.

Research on human judgement of contingencies using such paradigms has shown that humans are generally good at detecting contingencies from structured presentations: contingency ratings typically correlate highly with actual contingencies in cue-outcome presentations (8). However, some systematic biases have also been observed. For example, people's ratings of contingencies depend on how frequently the critical outcome is presented (as opposed to no-outcome). This has been termed the outcome density effect: a higher outcome density reliably leads to higher judgements of contingency (for the same true contingency level) (3,4,9).

Here, we report another systematic bias in contingency judgements, which we label the *negative contingency illusion*. A negative contingency illusion occurs when cue and outcome are truly unrelated,

but the contingency is judged and reported as *negative* – the cue is thought to be predictive of the *non-outcome*. Such inaccurate judgements could be important for human behavior and decision making if they lead individuals to conclude that an action or stimulus is protective when it has in fact no value.

Examples of possible negative contingency illusions appear in the data of a few studies of contingency learning (4,10,11), but the authors of these studies did not draw attention to or explore the effect, perhaps because the finding is not explained by standard associative accounts (12). Here, we report two experiments that demonstrate that a negative contingency illusion is reliably produced in participants under defined conditions: when outcome density is relatively low, and the contingency is either absent or modestly positive. In the context of illness, this implies that individuals may routinely judge certain situations or actions as protective, even though they have in actuality no protective value, or more concerningly, are in actuality modestly associated with greater risk of getting sick.

# Methods

All experiments were performed in accordance with relevant guidelines and regulations, including the Declaration of Helsinki.

**Experiment 1:** We followed the same general design as Crump and colleagues (4), with some modifications. While they (4) used simple shapes (circle and square) as cue and outcome events, and a relatively fast 100 ms frame rate to present all cue-outcome events in each streamed-trial, we aimed to ensure that their results were not idiosyncratic to those conditions, and modified both the visual depictions of cues and outcomes as well as frame rate. First, cues were pictures of foods and outcomes were cartoon depictions of allergic reactions. Second, we used two frame rates, presenting cue-outcome events for either 100 ms or 300 ms at a time. The slower frame rate makes the procedure more similar to standard contingency tasks that present event pairs one trial at a time. Similarly to Crump and colleagues (4), we had subjects rate the contingency between a cue and outcome in four conditions: ΔP zero/ outcome density low, ΔP zero/ outcome density high, ΔP medium/ outcome density low, and ΔP medium, outcome density high. The primary question was whether subjects would give negative contingency ratings for zero contingency/low outcome density conditions.

**Participants.** Ten students from Vanderbilt University took part. All subjects were compensated with research participation credit for a psychology course. All subjects had normal or corrected-to-normal vision, and provided written informed consent. All experimental protocols were approved by the Vanderbilt University Institutional Review Board.

**Apparatus & Stimuli.** The experiment was controlled by PCs running in-house LIVECODE software. Subjects sat at a viewing distance of approximately 60 cm from the computer screen. During each trial, subjects viewed a stream of 60 cue–outcome pairs. Each cue–outcome pair was presented in a frame (8.5 cm in height, 5.5 cm in width), displayed in the center of the screen in grey against a white background. There were four possible cue – outcome pairs, and examples of each of these pairs (e.g., A, B, C, D) are depicted in Figure 1A. Cues were always presented in a square (3.3 cm), centered at the

bottom of the frame, and outcome were always presented in a square (3.3 cm), centered at the top of the frame. Each streamed trial used a unique cue randomly selected from a set of 17 pictures of food items, and a unique outcome randomly selected from a set of four cartoon depictions of allergic reactions.

**Design & Procedure.** The experiment used a 2x2x2 within-subjects design involving contingency (noncontingent: $\Delta P=0$, vs. contingent: $\Delta P= .47$), outcome density (low: $P(O)=.27$ vs. high: $P(O)=.73$), and stream speed (Duration and inter-frame interval 100 ms vs. 300 ms). There were four streamed-trials in each condition, for a total of 32 trials. The stream speed factor was manipulated between two blocks of 16 trials, with order counterbalanced across subjects.

Each streamed-trial involved the rapid-serial-visual-presentation (RSVP) of 60 frames of cue-outcome pairs. The order of frames within each stream was randomly generated from one of the four contingency structures listed in Table 1, which defines the frequency of each cue-outcome event for each condition. These cell frequencies were adapted from (4) and modified so that each level of contingency had the same two levels of outcome density (in prior work the difference between outcome density levels was larger for the noncontingent than contingent conditions).

Within a stream, each frame was displayed for 100 (or 300) ms with a 100 (or 300) ms inter-frame interval (blank screen). In total, each stream lasted approximately either 12 seconds or 36 seconds. Following each streamed-trial, subjects gave contingency ratings using a continuous scroll-bar that participants could vary between a maximum negative value (– 100) and a maximum positive value (+100).

We used the same general procedure and instructions described in (4). Subjects first received the following general instructions:

“On each trial you will watch a stream of frames depicting a food item and an allergic reaction. The food item always appears on the bottom, and the allergic reaction always appears on the top. The presence or absence of the item represents eating or not eating the food. The presence or absence of the allergic reaction represents the presence or absence of the symptom. After each trial your task will be to rate the association between eating the food and developing an allergic reaction.”

Next, subjects viewed example streams containing a medium positive contingency ($\Delta P=.47$), zero contingency ($\Delta P=0$), and medium negative contingency ($\Delta P=-.467$). Each example was accompanied with definition of positive, zero, and negative contingencies. For brevity, we describe the instructions and definition for the positive contingency example. This format was modified appropriately for the descriptions of zero and negative contingency examples.

“A positive contingency means that eating the food predicts the allergic reaction. In other words, most of the time when the food item appears, the allergic reaction also appears; and, most of the time when the food does not appear, the allergic reaction does not appear. When you press the mouse button, you will be

given an example that demonstrates a positive association. You will notice that food usually appears with the allergic reaction. And, when the food is not presented, the allergic reaction will usually not occur.

You just viewed a positive association. In the real experiment, you will be asked to rate the strength of this association. You will be shown a rating bar, and you will need to click on the line to rate the strength of association. In the example that you just saw, the association was about +50, so you would want to click about here (the instructions display the rating scale, and an arrow points to the indicated position)."

Following the instruction phase, subjects viewed and gave contingency ratings for each of the streamed trials.

### Experiment 2

**Participants.** Twenty students from Brooklyn College took part. All subjects were compensated with course credit. All subjects had normal or corrected-to-normal vision, and provided written informed consent. All experimental protocols were approved by the Brooklyn College Institutional Review Board.

**Apparatus & Stimuli.** With the exception that the subjects completed the tasks on iMacs, the apparatus and stimuli were the same as Experiment 1.

**Design & Procedure.** The experiment used a 3x7 within-subjects design involving contingency ($\Delta P$=0, .13, .267), and outcome density (P(O)=.167, .267, .367, .467, .567, .667, and .767). Table 2 lists the cue-outcome event frequencies for each condition. Streamed trials from all 21 conditions were presented in each block, for a total of four blocks, and 84 total trials. Condition order was randomized within each block. Frame duration and inter-frame interval were both set to 100 ms. The same procedure and instructions from Experiment 1 were used for Experiment 2.

# Results

Experiment 1

In Experiment 1 we tested whether the negative contingency illusion is reliably observed across task parameters in a discrete-trials procedure, and whether it specifically arises when outcome density is low. On each trial, a series of cue-outcome pairings (cue-outcome; cue-no outcome; no cue-outcome; no cue-no outcome) was presented, followed by a rating of the contingent relationship that was observed between the cue and outcome (See Fig. 1A). In the Crump et al. (4) study in which there was evidence for a negative contingency illusion, the investigators used simple shapes (circle and square) as cue and outcome events, and a relatively fast 100 ms frame rate to present all cue-outcome events in each streamed-trial. To ensure that results were not idiosyncratic to those parameters, we modified both the visual depictions of cues and outcomes as well as frame rate. Cues were pictures of foods and outcomes were cartoon depictions of allergic reactions, which allowed for a more ecologically relevant prediction than simple geometric shapes. We used two frame rates in the experiment, presenting cue-outcome events for either 100 ms or 300 ms, with the slower frame rate corresponding more closely to exposure

times in standard contingency tasks that present event pairs one trial at a time. Subjects rated the contingency between a cue and outcome in four conditions: ΔP zero/ outcome density low, ΔP zero/ outcome density high, ΔP medium/ outcome density low, and ΔP medium, outcome density high. Our primary question was whether subjects would give negative contingency ratings for zero contingency/low outcome density conditions.

The mean contingency ratings in each condition for each subject were submitted to a 2x2x2 repeated measures ANOVA with contingency (noncontingent: ΔP=0, vs. contingent: ΔP= .47), outcome density (low: P(O)=.27 vs. high: P(O)=.73), and stream speed (Duration and inter-frame interval 100 ms vs. 300 ms) as factors. The means are displayed in Figure 1B.

We found a strong effect of contingency, $F(1, 9) = 37.69$, MSE =794.11, $p<.001$; and a moderate effect of outcome density, $F(1,9) = 5.08$, MSE =1560.29, $p=.05$; and no effect of stream speed. The pattern of contingency ratings reproduced the same pattern found by Crump and colleagues (4), with subjects giving higher ratings for contingent than noncontingent streams, and for streams with higher than lower outcome density.

Critically, we observed a strong negative contingency illusion. Specifically, subjects gave negative ratings for the low-outcome density, non-contingent streams (mean -26 (interval 100 ms) and -23 (interval 300 ms)). For both stream speeds of 100 and 300 (ms), one-sample t-tests (2-tailed) indicated these negative ratings were significantly different from 0, $t(9) = 4.61$, $p=.001$; and, $t(9) = 3.93$, $p=.003$).

Experiment 2

The purpose of Experiment 2 was to establish boundary conditions for the negative contingency illusion. Given the apparent robustness of the effect in Experiment 1, we were particularly interested in whether the negative contingency illusion would persist even when the programmed ΔP value increases from 0 to positive values. A second question addressed which levels of outcome density produce the strongest negative contingency illusion. To address these questions, we parametrically varied both ΔP and outcome density in small steps using the stream trials approach.

The mean contingency ratings in each condition for each subject were submitted to a 3x7 repeated measures ANOVA with contingency (ΔP=0, .13, .267), and outcome density (P(O)=.167, .267, .367, .467, .567, .667, and .767) as factors. The means for each condition are displayed in Figure 1C. As expected, there was a main effect of contingency, $F(2, 38) = 3.21$, MSE =478.12, $p=.05$; and a main effect of outcome density, $F(6,114) = 9.73$, MSE =750.45, $p<.0001$.

Figure 1C shows that mean ratings for the lowest outcome density condition (P(O)=.167) were all negative regardless of the ΔP condition. One-sample t-tests of the mean contingency rating against zero showed significant negative ratings in the ΔP=0 (mean rating = -19, $t(19)=3.81$, $p=.001$), ΔP = .13 (mean rating = -14, $t(19)=2.71$, $p=.014$), and ΔP = .267 conditions (mean rating = -13, $t(19)=2.79$, $p=.011$). Thus, even in conditions in which there was a positive association between stimulus and outcome, the low

outcome density produced a negative illusion of contingency. At the next lowest outcome density (P(O) = .267), subjects on average reported negative contingencies for both ΔP=0 and .13, but these were not significantly less than zero. None of the other contingency ratings across conditions were negative and significantly different from zero.

# Discussion

We show that a negative contingency illusion is reliably induced under defined conditions: when outcome density is relatively low, and the contingency is absent or modestly positive. Under such conditions, people perceive a cue-outcome contingency that is entirely absent as *negative*, misinterpreting that the cue makes the outcome less likely than chance, deeming it to have a 'protective' value. Strikingly, we show that this illusion persists even for moderately positive cue-outcome contingencies when outcome density is low. In such cases, despite the positive contingency between the cue and outcome, participants perceive the reverse relation and report a negative contingency. This is a potentially dangerous misjudgement if an individual misinterprets a cue or action as protective, when in actuality it is positively associated with an adverse outcome. We appear particularly vulnerable to such misjudgements when the actual prevalence of a problem is relatively rare, as is the case for many diseases.

Conceptually, a negative contingency indicates a 'protective factor', meaning that the presence of a cue makes the outcome less likely than chance. A negative contingency illusion could contribute to explaining how protective superstitions develop, over and above negative reinforcement. For disease *prevention*, a negative contingency illusion could lead to beliefs of positive benefits for scientifically unvalidated treatments that have no effect. For disease *treatment*, a negative contingency illusion might lead to placebo effects or belief in the benefits of treatments that have either no benefit or are even modestly deleterious. This has major public health and economic implications, given the wide number of products such as herbal supplements that are marketed without having been demonstrated to be beneficial through rigorous randomized clinical trials. Indeed, in the context of COVID-19 prevention, such a negative contingency illusion could contribute to belief in the benefits of previously unproven drug treatments.

An important question arises as to why humans would be vulnerable to such a potentially critical bias in judgement. We propose the findings may be explained by instance-based principles involved in encoding and retrieving event knowledge from memory (13). People may generate cue-outcome expectations by a process of cued-recall (14,15), whereby thinking about a cue causes retrieval of prior traces with the cue, along with prior paired outcomes contained in the traces. This process of cued recall allows contingency judgements through which cued-recall is used to generate both an expectation for an outcome given a cue was present (P(O|C)), and an expectation for an outcome given a cue was absent (P(O|~C)). In low density outcome situations, attending to just (P(O|C) will automatically produce more instances of the outcome not occurring following the cue than outcomes. For instance in Study 1, probing a memory for those events with cue present information would cause retrieval of more no outcome traces (C ~ O = 22)

than outcome traces (CO = 8), and on this basis alone, the cue might appear to be protective against the outcome, and support a negative contingency judgement. This is why accurate contingency judgements require incorporation of information on (P(O|~C). However, differences in the encoding and retrieval of (P(O|~C) and (P(O|C) may lead to a critical bias. Specifically, because memory retrieval is guided by cue-driven similarity, the content of what is retrieved depends on the content of the cue probing memory. In the context of a contingency judgement, probing memory for cue-present information, like a specific "strawberry", will result in selective retrieval of events (and their outcomes) in memory that are similar to the specific strawberry in mind. However, it is less clear how people probe their memory for cue-absent information: there are multiple options, such as somehow imagining the absence, omission or opposite of a cue, or perhaps recalling multiple different irrelevant cues, among other strategies. The implication is that memory-based outcome expectations based on cue-absent information will be systematically noisier and of weaker confidence than for cue-present information. The weaker quality of such information may therefore lead to it receiving lesser weighting when attempting to integrate it with cue-present information, thus biasing the contingency judgement toward memory-based expectations driven by cue-present information. Without properly calibrating for, or weighting the rate of outcomes when the cue is absent, even a cue that modestly positively predicts an outcome can appear protective in low outcome density situations in which the outcome occurs less often than its absence.

Understanding of the negative contingency illusion may in the future benefit from a more formal memory-based computational model of contingency judgements, including for instance D. L. Hintzman's (14,16) MINERVA 2 computational model of instance-based memory, which has notably been extended into the domain of decision-making (e.g., Dougherty and colleagues' (17), MINERVA-DM for judgements of likelihoods and biases), and the domain of associative learning (18). In our view, it offers a promising and potentially integrative framework for examining contingency judgement phenomena in general.

The negative contingency illusion may be particularly important in the current climate in which there is a desire to rapidly find treatments that might prevent a disease like COVID-19. In the absence of randomized control trials, drugs (take hydroxychloroquine for example) may appear to have preventative value based on just such an illusion. Imagine the simple scenario, a person knows of 20 people who took the medication and only two of them went on to develop COVID-19. Even if this is actually slightly worse than the rate of disease among those who did not take the medication, the presence of the negative contingency illusion makes it likely that many people will mistakenly conclude that the drug has at least a modest benefit.

# Declarations

**Acknowledgments**.

**Funding:** R01 MH074567

**Author contributions:** RMP, MJCC, DHZ designed the study; MJCC carried out and analyzed the study; RMP, MJCC, DHZ wrote the manuscript;

**Competing interests:** Authors declare no competing interests.

**Data and materials availability:** All data is available from the corresponding author upon request.

# References


1. L. G. Allan, H. M. Jenkins, The effect of representations of binary variables on judgement of influence. *Learning and Motivation*. **14**, 381–405 (1983).

2. L. G. Allan, Human contingency judgements: Rule based or associative? *Psychological Bulletin*. **114**, 435–448 (1993).

3. L. G. Allan, S. Siegel, J. M. Tangen, A signal detection analysis of contingency data. *Animal Learning & Behavior*. **33**, 250–263 (2005).

4. L. B. Alloy, L. Y. Abramson, Judgement of contingency in depressed and nondepressed students: Sadder but wiser? *Journal of Experimental Psychology: General*. **108**, 441–485 (1979).

5. M. J. C. Crump, S. D. Hannah, L. G. Allan, L. K. Hord, Contingency judgements on the fly. *The Quarterly Journal of Experimental Psychology*. **60**, 753–761 (2007).

6. A. Dickinson, D. Shanks, J. Evenden, Judgement of act-outcome contingency: The role of selective attribution. *The Quarterly Journal of Experimental Psychology Section A*. **36**, 29–50 (1984).

7. M. R. P. Dougherty, C. F. Gettys, E. E. Ogden, MINERVA-DM: A memory processes model for judgements of likelihood. *Psychological Review*. **106**, 180–209 (1999).

8. D. L. Hintzman, MINERVA 2: A simulation model of human memory. *Behavior Research Methods, Instruments, & Computers*. **16**, 96–101 (1984).

9. D. L. Hintzman, Judgements of frequency and recognition memory in a multiple-trace memory model. *Psychological Review*. **95**, 528–551 (1988).

10. L. L. Jacoby, L. R. Brooks, in *Psychology of Learning and Motivation*, G. H. Bower, Ed. (Academic Press, 1984), vol. 18, pp. 1–47.

11. R. K. Jamieson, S. D. Hannah, M. J. C. Crump, A memory-based account of retrospective revaluation. *Canadian Journal of Experimental Psychology/Revue canadienne de psychologie expérimentale*. **64**, 153–164 (2010).

12. M. E. Le Pelley, S. M. Oakeshott, I. P. L. McLaren, Blocking and Unblocking in Human Causal Learning. *Journal of Experimental Psychology: Animal Behavior Processes*. **31**, 56–70 (2005).

13. S. C. Musca, M. A. Vadillo, F. Blanco, H. Matute, The role of cue information in the outcome-density effect: evidence from neural network simulations and a causal learning experiment. *Connection Science*.

**22**, 177–192 (2010).

14. R. A. Rescorla, A. R. Wagner, in *Classical Conditioning II: Current Research and Theory*, Black AH, Prokasy WF, Eds. (New York: Appleton Century Crofts, 1972), pp. 64–99.

15. R. Semon, *Mnemic Psychology* (George Allen And Unwin Limited, 1923; http://archive.org/details/mnemicpsychology032279mbp).

16. D. R. Shanks, Continuous monitoring of human contingency judgement across trials. *Memory & Cognition*. **13**, 158–167 (1985).

17. D. R. Shanks, Selective attribution and the judgement of causality. *Learning and Motivation*. **17**, 311–334 (1986).

18. D. R. Shanks, A. Dickinson, in *Psychology of Learning and Motivation*, G. H. Bower, Ed. (Academic Press, 1988; http://www.sciencedirect.com/science/article/pii/S0079742108600304), vol. 21, pp. 229–261.

# Tables

Table 1.

Frequencies of cue-outcome pairings, outcome density, and ΔP for conditions in Experiment 1

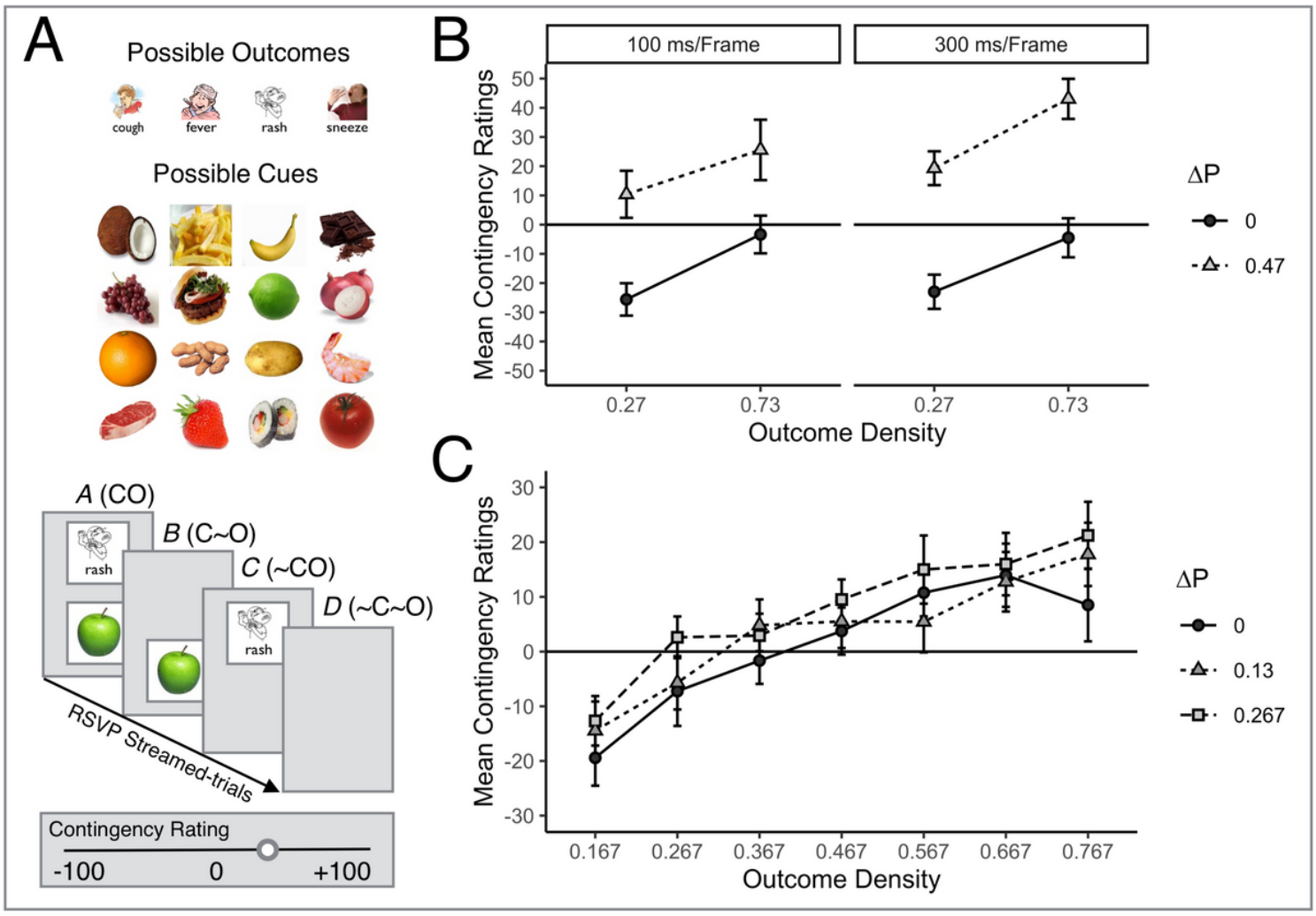
| A (CO) | B (C~O) | C (~CO) | D (~C~O) | P(O) | ΔP |
|---|---|---|---|---|---|
| 8 | 22 | 8 | 22 | 0.27 | 0.00 |
| 22 | 8 | 22 | 8 | 0.73 | 0.00 |
| 15 | 15 | 1 | 29 | 0.27 | 0.47 |
| 29 | 1 | 15 | 15 | 0.73 | 0.47 |

Table 2.

Frequencies of cue-outcome pairings, outcome density, and ΔP for conditions in Experiment 2.

| A (CO) | B (C~O) | C (~CO) | D (~C~O) | P(O) | ΔP |
|---|---|---|---|---|---|
| 5 | 25 | 5 | 25 | 0.167 | 0 |
| 7 | 23 | 3 | 27 | 0.167 | 0.13 |
| 9 | 21 | 1 | 29 | 0.167 | 0.267 |
| 8 | 22 | 8 | 22 | 0.267 | 0 |
| 10 | 20 | 6 | 24 | 0.267 | 0.13 |
| 12 | 18 | 4 | 26 | 0.267 | 0.267 |
| 11 | 19 | 11 | 19 | 0.367 | 0 |
| 13 | 17 | 9 | 21 | 0.367 | 0.13 |
| 15 | 15 | 7 | 23 | 0.367 | 0.267 |
| 14 | 16 | 14 | 16 | 0.467 | 0 |
| 16 | 14 | 12 | 18 | 0.467 | 0.13 |
| 18 | 12 | 10 | 20 | 0.467 | 0.267 |
| 17 | 13 | 17 | 13 | 0.567 | 0 |
| 19 | 11 | 15 | 15 | 0.567 | 0.13 |
| 21 | 9 | 13 | 17 | 0.567 | 0.267 |
| 20 | 10 | 20 | 10 | 0.667 | 0 |
| 22 | 8 | 18 | 12 | 0.667 | 0.13 |
| 24 | 6 | 16 | 14 | 0.667 | 0.267 |
| 23 | 7 | 23 | 7 | 0.767 | 0 |
| 25 | 5 | 21 | 9 | 0.767 | 0.13 |
| 27 | 3 | 19 | 11 | 0.767 | 0.267 |

# Figures




## Figure 1

The streamed-trials contingency judgment task and results from Experiment 1 and 2. (A) The top panel shows the stimulus set of four possible outcomes and sixteen possible cues. One cue and outcome were randomly selected for each streamed trial. The lower panel shows all possible example frames from a streamed trial. Cues were presented on the bottom of each frame, and outcomes were presented on the top. Participants gave a contingency rating on a slider at the end of each streamed trial. (B) Mean contingency ratings for Experiment 1 as a function of streamed trial frame rate, ΔP, and outcome density. Participants gave negative ratings for low outcome density trials (.27) with no programmed relationship (ΔP = 0). (C) Mean contingency ratings for Experiment 2 as a function of ΔP and Outcome density. Participants gave negative ratings for low outcome density trials (.167) with no programmed relationship (ΔP=0), and for trials with a weak positive relationship (ΔP=.13, .267).