Psychophysics of Associative Learning: Quantitative Properties of Subjective Contingency

Susana Maia, Francoise Lefevre, Jeremie Jozefowiez

▶ To cite this version:

Running head: PSYCHOPHYSICS OF ASSOCIATIVE LEARNING

Psychophysics of Associative Learning: Quantitative Properties of Subjective Contingency

Susana Maia

Universidade do Minho

Françoise Lefèvre

Université de Lille & Université Catholique de Louvain

Jérémie Jozefowiez

Université de Lille & Universidade do Minho

Correspondence:

Dr. Jérémie Jozefowiez

Laboratory of Cognitive and Affective Sciences (SCALab)

Université de Lille

Campus de Lille 3

Domaine Universitaire du Pont de Bois

B.P. 60149

59653 Villeneuve d'Ascq Cedex

France

Email: jeremie.jozefowiez@univ-lille3.fr

Phone: +33-320-041-6866

Fax: +33-320-041-6036

Submitted: July 2017

Abstract

Allan and collaborators recently proposed to apply signal detection theory (SDT) to the analysis of contingency judgment tasks. When exposed to a flow of stimuli, participants are asked to judge whether there is a contingent relation between a cue and an outcome, that is, whether the subjective cue-outcome contingency exceeds a decision threshold. In this context, we tested the following hypotheses regarding the relation between objective and subjective cue-outcome contingency: (a) the underlying distributions of subjective cue-outcome contingency are Gaussian; (b) the mean distribution of subjective contingency is a linear function of objective cue-outcome contingency; (c) the variance in the distribution of subjective contingency is constant. The hypotheses were tested by combining a streamed-trial contingency assessment task with a confidence rating procedure. Participants were exposed to rapid flows of stimuli at the end of which they had to judge whether an outcome was more (Experiment 1) or less (Experiment 2) likely to appear following a cue and how sure they were of their judgment. We found that while hypothesis (a) seems reasonable, hypotheses (b) and (c) were not. Regarding hypothesis (b), participants were more sensitive to positive than to negative contingencies. Regarding hypothesis (c), the perceived cue-outcome contingency became more variable when the contingency became more positive or negative, but only to a slight extent.

Keywords: Associative learning, contingency judgment, psychophysics, signal detection theory, streamed-trial procedure.

Psychophysics of Associative Learning: Quantitative Properties of Subjective Contingency

In a contingency judgment task, participants exposed to flows of stimuli are asked to assess whether a specific stimulus (the outcome) is more or less likely to appear in the presence of another one (the cue). Contingency judgments in such tasks are a function of the objective cue-outcome contingency, as measured by the ΔP index, which is the difference between the probability of the outcome in the presence of the cue and its probability in the absence of the cue (Allan, 1993; Schanks, 2007; Rescorla, 1967). When $\Delta P = 0$ (null contingency), participants are less likely to expect the outcome in the presence of the cue than if ΔP is positive (positive contingency), but are otherwise more likely to do so than if ΔP is negative (negative contingency). Theories of associative learning explain such results by the building up of some internal quantity, e.g. the strength of the association between the cue and the outcome (i.e. Rescorla & Wagner, 1972; Mackintosh, 1975; Miller & Matzel, 1988; Pearce & Hall, 1980; Wagner, 1980), even though non-associative accounts have also been proposed (i.e. Gallistel & Gibbon, 2000). We refer to this quantity as "subjective contingency" throughout this paper.

Besides subjective contingency, contingency judgments are also affected by other variables, notably response bias (Allan, Hannah, Crump, & Siegel, 2008; Perales, Catena, Schanks, & Gonzales, 2005), making them an imperfect indicator of learning. For this reason, Allan and collaborators (Allan et al., 2008; Alan, Siegel, & Tangen, 2005; Siegel, Crump & Allan, 2009. See also Laux, Goedert, & Markman, 2010 and Perales et al., 2005) proposed applying signal detection theory (SDT) (Wickens, 2002) to contingency judgment tasks. In this view, subjective contingency would be a random variable x. If its value exceeds a threshold λ, the subject reports perceiving a contingency between the cue and the outcome. Variables such as response bias are assumed to affect only λ. As SDT can be used to

disentangle the contributions of *x* and λ to performance, it provides a bias-free assessment of learning (Wickens, 2002).

In this context, the present paper aims to test three hypotheses regarding the quantitative properties of *x*: (a) subjective contingency *x* is drawn from a Gaussian distribution with mean μ and standard deviation σ; (b) μ is a linear function of ΔP; (c) σ is constant. Hypotheses (a) and (c) are often assumed by default when applying SDT to data because they simplify the computation needed to retrieve the parameters of the SDT model. Indeed, they subsume the thinking of Allan et al. Hypothesis (b) stems from the Rescorla-Wagner (1972) model in which the associative strength between the cue and the outcome converges on ΔP at the asymptote (Danks, 2003).

When the variance of the distribution is allowed to vary in a Gaussian SDT model, the decision criterion λ has to be manipulated in order to retrieve the parameters of the model from the data (see Wickens, 2002 for the basic theory). Though the criterion can be manipulated explicitly, e.g. by using differential payoffs for each response, it is more often manipulated implicitly by using the confidence rating technique. Participants are asked how confident they are about their choice after categorizing a stimulus. The idea is that each confidence rating corresponds to its own decision criterion. A change in λ can then be simulated through the confidence rating data (see Wickens, 2002 for more details).

The present study applies the confidence rating procedure to a contingency assessment task. Participants were exposed to various cue-outcome contingencies ranging from strongly negative to strongly positive. After each ΔP condition, they had to decide whether the outcome was contingent upon the cue ("yes" vs. "no" responses) and how sure they were of their judgment ("very sure" vs. "sure" vs. "not sure" responses). Because the confidence rating technique requires accurate estimates of the probability of responding for each combination of response category, confidence level and stimulus condition, each ΔP

condition has to be presented as many times as possible. This led us to use a streamed-trial procedure in which cues and outcomes lasted only for a few hundreds of milliseconds, allowing for the maximal exposure of the participants to the stimulus conditions in a relatively short time. This procedure was designed by Allan et al. (Crump, Hannah, Allan, & Hord, 2007; Hannah, Crump, Allan, & Siegel, 2009; Siegel et al., 2009) specifically to allow the SDT analysis of contingency learning.

<div align="center">Experiment 1</div>

<u>Method</u>

Participants and apparatus

Seventeen participants (8 males, 9 females) aged from 18 to 41 years old (mean: 25.88) were recruited for this experiment. All experimental tasks were conducted on an IBM compatible notebook PC. All stimuli were displayed on a 15.6-inch HD TFT monitor screen (pixel resolution: 1366 x 768). Participants sat about 50 cm from the screen. A customized program written in Visual Basics controlled all experimental events and recorded the results.

Procedure

The experiment was composed of two experimental sessions. The first session started with three pretesting phases the goal of which was to familiarize participants with the procedure and to show them examples of strong positive and negative contingencies. This was followed by the first testing phase. The second session started right away with the second testing phase and was identical to the previous testing phase. The testing was split in two phases to avoid having an experimental session which would be too long for the participants.

Pretesting I: After reading the instructions informing them of the task they had to perform, the participants were presented with two successive streams of stimuli. The first stream started with a 1-s black screen with a fixation cross in its center. The cross remained visible until the end of the stream. This was followed by 20 trials. Each trial lasted for 200 ms

and comprised two phases: (a) the cue phase lasting 100 ms in which a yellow triangle

(stimulus A) was always displayed in the upper right-hand corner of the screen, marking the

beginning of a trial. On some trials, a green circle (target cue X) was displayed in the upper

left-hand corner of the screen; (b) the outcome phase lasting 100 ms. In some trials, a red

square (outcome O) was presented in the lower central part of the screen. All stimuli

measured 501 by 499 pixels. Trials were separated by a 100-ms inter-trial interval (ITI)

during which only the fixation cross was displayed.

At the end of a stream, a dialog box appeared in the center of the screen with the

question: "Is the red square more likely to appear following the green circle?". The

participants used the mouse to answer (the pointer always appeared centrally between the two

response buttons) by clicking either on the "Yes" or on the "No" button displayed below the

question. Once they clicked on one of the buttons, the dialog box disappeared and another

one appeared with the question "How sure are you of your decision?" with three buttons,

"Not sure", "Sure", and "Very sure", appearing below the question. Once the participant

answered this question, the dialog box disappeared and a new stream started, except if this

was the end of a phase, in which case a debriefing screen was shown.

Overall, a stream was composed of four types of trials: (a) AX+ trials in which both A

and X were shown during the cue phase, the outcome being shown during the outcome phase;

(b) AX- trials in which both A and X were shown during the cue phase, the outcome not

being shown during the outcome phase; (c) A+ trials in which only A was shown during the

cue phase, the outcome being shown during the outcome phase; (d) A- trials in which only A

was shown during the cue phase, the not being shown during the outcome phase.

Manipulating the proportion of these four types of trials within a stream induced different ΔP

values between the target cue X and the outcome.

During the first phase of training, each stream was composed of 10 AX+ and 10 A-

trials ($\Delta P = 1$). The order of presentation of the trials within a stream was determined randomly. Once the participants had been through the two stimulus streams, they were presented with a debriefing instruction explaining to them why they should have answered "Yes" to the contingency question.

Pretesting II: The second phase of pretesting started immediately after the participants had completed the first phase of training. It was identical to the previous phase of pretesting except that a stream was now composed of 10 AX- trials and 10 A+ trials ($\Delta P = -1$). At the end of the phase, a debriefing instruction explained to the participants why they should have answered "No" to the contingency question.

Pretesting III: The third phase of pretesting started immediately after completion of the second phase of training. Participants were exposed to two streams. One of them was composed of 10 AX+ trials and 10 A- trials ($\Delta P = 1$), while the other was composed of 10 AX- trials and 10 A+ trials ($\Delta P = -1$). The order of presentation of the streams was determined randomly. The participants were expected to answer "Yes" when asked about the relation between X and O following the $\Delta P = 1$ stream, and to answer "No" following the $\Delta P = -1$ stream. If they failed to do so for at least one stream, the two streams were presented over again until they reached this learning criterion or until an upper limit of 10 repetitions was reached. At the end of the phase, a debriefing instruction informed them that testing was about to start. The instruction also indicated that discriminating between the positive and negative contingencies would be more difficult during testing than during training and that the contingencies between the cue and the outcome would vary randomly from stream to stream.

First testing phase: Testing began as soon as the pretesting was over. Participants were exposed to streams with a ΔP between the cue and the outcome of either -0.8, -0.4, 0, 0.4 or 0.8. The composition of the trials for each of these streams is shown in Table 1. Each

ΔP condition was presented 30 times to a participant. The order of presentation of each ΔP condition as well as the order of each trial type within a stream was determined randomly.

As shown in Table 1, both components of ΔP (P(O|AX), the probability of the outcome following the cue, and P(O|A), the probability of the outcome to be presented without the cue), varied together from one ΔP condition to the next. This was done to ensure that P(O), the overall probability of the outcome (sometimes called the outcome density) remained constant across streams (Table 1). Allan et al. (2005) argued that the outcome density affects the decision criterion λ. If this is correct and if the outcome density had been different between streams, we could not have used the confidence rating technique as it assumes that the decision criteria remain constant across stimulus conditions. Moreover, if the outcome density had changed between streams, the participants could have used that information to discriminate between ΔP conditions. By keeping it constant, we made sure that they engaged in a form of associative learning. The drawback is that it is impossible to know whether the performance of a participant relies on P(O|AX), on P(O|A) or on a combination of the two. As this was never the goal of this study, we deemed that the advantages of keeping P(O) constant outweighed its disadvantages. Note that other variables which have been suggested as a potential determinants of the decision criterion λ (the frequency of the cue within a stream: Perales et al., 2005; the contingency between the alternative cue A and the outcome: Laux et al., 2010; Hannah & Allan, 2011 Siegel et al., 2009) were also kept constant across streams.

Second testing phase: The second testing phase was identical to the first one. It took place later the same day or at most the day after, depending on the availability of the participants.

Data analysis

For each participant, the 15 parameters of the SDT model (the mean and standard

deviation of the five subjective contingency distributions corresponding to the five ΔP values plus the Yes/No response threshold and the four confidence rating criteria) were retrieved from the confidence rating data using an iterative maximum-likelihood algorithm (MLA), as recommended by Wickens (2002) and Macmillan & Creelman (2005). We used the MLA implemented by Harvey's (2013) RscorePlus program. The program "uses singular value decomposition, combined with a variation of the Marquardt method for nonlinear least-squares regression (Marquardt, 1963; Press, Teukolsky, Vetterling, & Flannery, 2002, 2007), to find the maximum likelihood fit of the multiple distribution, a variable-criterion signal detection model to confidence rating-scale data" (Harvey, 2013). Parameters of an SDT model are always expressed relative to a reference condition the mean of which is set to 0 and standard deviation to 1. We chose the $\Delta P = 0$ condition to be that reference. As ratings with a frequency of 0 could prevent the MLA from converging on a solution, a loglinear correction was used, where 1/m (m being the number of possible responses given by the participant, in our case, m = 6) is added to all the data (See Hautus & Lee, 1998).

Based on the likelihood computed by the program that the data could have been generated by the SDT model, a χ^2 (with 12 degrees of freedom) was computed to quantify the fit between the model and the data. A significant χ^2 at the conventional threshold of 0.05 was interpreted as indicating a potential discrepancy between the model and the data.

In the following, we focus on the estimates of the mean and standard deviation of the subjective contingency. We do not report the decision and confidence thresholds as they were of no particular interest to us.

The analysis below fits a second-order polynomial equation to the function mapping ΔP onto either the mean of the subjective contingency distribution or its standard deviation. To allow for a meaningful interpretation of the parameters of the polynomial equation, each ΔP value was divided by 0.8, creating new recoded ΔP values ranging from -1 to 1. These

recoded values were used to perform the polynomial regression. This recoding made the quadratic and the linear terms of the polynomial equation independent from each other. In these conditions, the linear term corresponds to the slope of the least-mean square linear equation that could have been fitted to the data while the quadratic term reflect the deviation from that straight line (Pagès, 2010).

All parameters based on an averaging of individual performance are reported along with their 95% confidence interval (CI), computed using Student's t distribution. All numbers are rounded to the nearest second decimal. Averages for the percentage of variance explained by a regression (r^2) and their 95% CI were computed based on the Fisher transform of the individual r^2. The inverse of the Fisher transform was then used to express them in terms of percentage of variance explained.

Results

Psychometric curves

The average proportion of trials in which a positive contingency was detected is shown in Figure 1 as a function of ΔP. It increased monotonically with ΔP according to a sigmoid-like function.

Is the subjective contingency drawn from a Gaussian distribution?

The χ^2 measuring the fit between the SDT model and the data was significant for 4 participants. The average proportion of trials on which a positive contingency was detected as a function of ΔP is shown for these 4 participants in Figure 2. Each curve and associated set of points show what the psychometric function would have looked like if the response criterion had been located at the corresponding confidence criterion.

P8 failed to discriminate between the various contingencies condition. For P1, the significant χ^2 seems to have been caused by the model under-predicting the probability to detect a positive contingency for the ΔP = 0.8 condition on two of the response criteria (No-

not sure vs. Yes-not sure and Yes-not sure vs. Yes-sure). Otherwise, the predictions of the model seem quite accurate so we decided to keep this participant in the analysis. There are clear quantitative discrepancies between the model and the data for participants P12 and P14, in the case of one response criterion for P12 (No-not sure vs. Yes-not sure) and of two in the case of P14 (No-not sure vs. Yes-not sure and Yes-not sure vs. Yes-sure). On the other hand, the deviations from the model from the data are not systematic and the model captures the overall trend in the data quite well, especially in the case of P12. Hence, in this case, we decided to reject P14 but to include P12 in the analysis.

Is the mean of the subjective contingency distribution a linear function of ΔP?

Figure 3 shows the mean of the subjective contingency distribution as a function of the recoded ΔP averaged across the 15 participants included in the analysis. Also shown is the best-fitting second-degree polynomial ($y = 0.30x^2 + 0.99x + 0.01, r^2 = 0.99$). As indicated by its quadratic term, there is a slight yet undeniable curvature in the function.

It is not an artefact of averaging as it can also be observed in the individual data. Table 2 shows for each participant the parameter of the second-degree polynomial equation $y = ax^2 + bx + c$ fitted to the function mapping the (recoded) ΔP onto the mean of the subjective contingency distribution, along with the resulting percentage of variance explained by the regression (r^2). The fit of the second-degree polynomial is exceptionally good. With the exception of P7, r^2 never falls below 0.95 while the average r^2 is equal to 0.99, 95% CI [0.97, 1.00]. Otherwise, the conclusions that can be drawn from Table 2 are the same as the one that can be drawn from Figure 3: The quadratic term of the polynomial equation is positive for all but one participant and the center of the distribution of quadratic terms is clearly above 0 (0.26, 95% CI [0.14, 0.38]).

Is the standard deviation of the subjective contingency distribution constant across ΔP values?

Figure 4 shows the standard deviation of the subjective contingency distribution as a function of (recoded) ΔP averaged across the 15 participants included in the analysis. Also shown is the best-fitting second-order polynomial ($y = 0.12x^2 + 0.02x + 1.00, r^2 = 0.92$). While the value of the standard deviation for the reference ΔP = 0 condition remains within the 95% CI of most of the other standard deviations, it falls outside of it for ΔP = -0.8, suggesting that variability in that condition was higher than in the reference condition. The non-zero quadratic term of the polynomial equation points in the same direction.

Table III shows the parameters and percentage of variance explained by the polynomial equation $y = ax^2 + bx + c$ which, for each participant, was fitted to the function mapping the recoded ΔP onto the standard deviation of the subjective contingency distribution. The average proportion of variance explained by the regression remains good (0.72, 95% CI [0.51, 0.85]), though clearly below the one observed for the means (Table II), with a wider range of variation across participants. The proportion of variance explained ranges from 0.04 to 0.98. At the qualitative level, the participants vary also much more than they did in the case of the mean. Almost all participants showed some quadratic and linear trends in Table 2. In Table 3, some participants show both trends (for instance, P5), one of them but not the other (P1: linear trend but no quadratic trend; P2: quadratic trend but no linear trend) or none of them (P10). Overall, however, the data point in the same direction as Figure 4. There seems to be a quadratic trend, though possibly a small one (average quadratic term: 0.12, 95% CI [0.02, 0.22]) and no linear trend (average linear term: 0.03, 95% CI [-0.05,0.10]).

Discussion

The aim of this study was to test the three hypotheses regarding the application of SDT to contingency assessment tasks: (a) the subjective contingency is drawn from a Gaussian distribution; (b) the mean of the subjective contingency distribution is a linear

function of ΔP; (c) its standard deviation is constant. Based on a χ^2 adjustment test, the SDT model provides a good account of the data of 13 out of 17 participants and provides at least a good qualitative account of 3 of the remaining participants (the fourth one failed to discriminate between the conditions). Overall, there is no strong evidence that hypothesis (a) is false.

Concerning hypothesis (b), the relation between ΔP and the mean of the subjective contingency distribution is clearly not linear. It is best fitted by a second-degree polynomial with a positive quadratic term, a fact that can be seen both in the average data and in the individual performances. From a psychological point of view, the interpretation of this positive curvature is that the participants were more sensitive to changes in the positive contingencies than in the negative ones.

Finally, concerning hypothesis (c), the data suggest the existence of a quadratic trend in the function mapping ΔP onto the standard deviation of the subjective contingency distribution. The variability of the perceived cue-outcome contingency increases as it becomes more positive or negative, though one could reasonably argue that the data provide only weak support for this conclusion.

Experiment 2

The higher sensitivity to the positive contingencies observed in Experiment 1 could be due to the fact that we asked the participants whether the outcome was *more* frequent following the cue. This could have biased them toward detecting positive contingencies (see Vadillo, Miller & Matute, 2005 for evidence of the effect of the contingency question). Hence, we carried out a new experiment aimed at assessing the impact of the contingency question. Experiment 2 was identical to Experiment 1 except that the participants were asked to judge whether the outcome was *less* frequent following the cue. If the higher sensitivity of the participants to positive contingencies observed in Experiment 1 was due to the

contingency question, we should observe a higher sensitivity to negative contingencies in Experiment 2. If, on the other hand, the higher sensitivity to positive contingencies observed in Experiment 1 was a feature of the human contingency assessment system, we should observe it again in Experiment 2.

Method

Participants and apparatus

Seventeen participants (2 males, 15 females) aged from 19 to 29 years old (mean: 22.00) were recruited for this experiment. The apparatus was identical to the one used in Experiment 1.

Procedure and data analysis

The procedure was identical to the one used in Experiment 1, except that at the end of a stream, participants were asked whether the outcome was *less* frequent following the target cue X. The instructions were modified accordingly. Data were analyzed the same way as in Experiment 1, except that what was the "yes" response in the current experiment was treated as the "no" response in the previous one and vice-versa.

Results

Psychometric curves

The average proportion of trials in which a negative contingency was detected is shown in Figure 5 as a function of ΔP. It decreases monotonically with ΔP according to a sigmoid-like function.

Is the subjective contingency drawn from a Gaussian distribution?

The χ^2 measuring the fit of the SDT model to the data was not significant for any participants. Nevertheless, we decided to exclude one participant (P2) from the analysis. In the $\Delta P = 0.8$ condition, this participant always answered "Yes – very sure", leading to an estimate of 15.98 for the mean of the $\Delta P = 0.8$ distribution. This value was so extreme

compared to the others (the maximum value for that mean in Experiment 1 is 2.84 while the largest value in Experiment 2 besides P2 was 2.28) that its inclusion would have distorted the averages.

Is the mean subjective contingency a linear function of ΔP?

Figure 6 shows the mean of the subjective contingency distribution as a function of the recoded ΔP averaged across the 16 participants included in the analysis. The best-fitting second-degree polynomial ($y = 0.30x^2 + 0.74x + 0.03, r^2 = 1.00$) is also shown. The function has a clearly positive curve.

The individual data show the same finding. Table 4 shows the parameters and r^2 values for the second-degree polynomial equation $y = ax^2 + bx + c$ used for each participant to predict the mean of the subjective contingency distribution on the basis of the (recoded) ΔP values. The polynomial equation accounts on average for a high proportion of the variance (0.97, 95% CI [0.93, 0.99]). Its quadratic parameter is positive for all but one participant (mean: 0.30, 95% CI [0.19, 0.41]).

Is the standard deviation of the subjective contingency distribution constant across ΔP values?

Figure 7 shows the standard deviation of the subjective contingency distributions as a function of the recoded ΔP averaged across the 16 participants included in the analysis. The best-fitting second-degree polynomial ($y = 0.09x^2 - 0.04x + 1.00, r^2 = 1.00$) is also shown. The curve shows a weak quadratic trend though, based on the 95% CI, only the standard deviation for ΔP = -0.8 is significantly different from that of the reference ΔP = 0 condition.

The same conclusion can be drawn from the individual data. Table 5 shows the parameters and r^2 values for the second-degree polynomial equation $y = ax^2 + bx + c$ used for each participant to predict the standard deviation of the subjective contingency

distribution on the basis of the (recoded) ΔP values. The polynomial equation accounts on average for a large proportion of the variance (0.56, 95% CI [0.30, 0.75]) but there is a wide range of variations between participants, with the proportion of variance ranging from 0.05 to 0.98. The individual differences in the pattern of variation of the standard deviation as a function of ΔP are as large as in Experiment 1, but taken as a whole, the data point to the existence of a small quadratic term in the relation existing between ΔP and the standard deviation of the subjective contingency distribution (mean quadratic term: 0.11, 95% CI [0.02, 0.14]), while ruling out a linear trend (mean linear term: 0.04, 95% CI [-0.08, 0.01]).

Discussion

The rationale behind Experiment 2 was that if the positive curvature observed in Experiment 1 in the function mapping ΔP onto the mean of the subjective contingency distribution was caused by our asking the participants whether the outcome was *more* frequent following the cue, the function should display a negative curvature in Experiment 2 in which we asked the participants whether the outcome was *less* frequent following the cue. The answer provided by the data is unambiguous. Whether the conclusions are based on the average means (Figure 6) or on the individual polynomial fit (Table 4), not only is the function mapping ΔP onto the mean of the subjective contingency distribution positively curved, but the degree of curvature seems identical to that observed in Experiment 1. The quadratic term of the second-degree polynomial fitted to the average data in Figure 3 is 0.29 vs. 0.30 in Figure 6. The average quadratic term is 0.30, 95% CI [0.19, 0.40] in Table 2 vs. 0.30, 95% CI [0.19, 0.41] in Table 4. In conclusion, Experiment 2 replicates our finding from Experiment 1, i.e. the participants were less sensitive to the negative contingencies than to the positive ones.

The results of Experiment 2 are strikingly like those of Experiment 1. The 95% CI of all the polynomial parameters and even of the r^2 values in Table 2 and 3 fall well within the

range of their counterparts in Table 4 and 5. Comparing Figure 3 to Figure 6 and Figure 4 to Figure 7, one can see that the parameters of the polynomial fitted to the data are also very similar. The linear term of the polynomial equation in Figure 6 is lower than that in Figure 3 but the comparison of the 95% CI in Table 2 and 4, to which more attention should be paid because of the strong link between ΔP and the mean of the subjective contingency distribution, does not support this conclusion. Hence, the change in the contingency question did not have any effect, so for practical purposes Experiment 2 can be considered as an independent replication of Experiment 1. This means that the results of both experiments can be compared to reach more definitive conclusions.

On this basis, it seems that the following conclusions can be reached with some level of certainty. The first is obvious: the participants were sensitive to the cue-outcome contingency, as indicated by Figures 1 and 5, Figures 3 and 4 and the positive value of the linear terms of the polynomial regression in Tables 2 and 4 (average linear parameter over Tables 2 and 4: 0.86, 95% CI [0.64, 1.07]).

Second, the hypothesis that the subjective contingency is drawn from a Gaussian distribution seems highly plausible given that, out of the 34 participants in both Experiments 1 and 2, a case could be made for a failure of the SDT model in only 5 participants.

Third, the mean of the subjective contingency distribution varies much more systematically with ΔP than its standard deviation. ΔP is almost the sole contributor to the variance of the mean (average r^2 over Table 2 and 4: 0.98, 95% CI [0.96, 0.99]). ΔP remains a main contributor to the variance of the standard deviation (average r^2 over Tables 3 and 5: 0.64, 95% CI [0.49, 0.76]) but other sources of variations clearly play a role.

The fourth conclusion is the one discussed above concerning the positive curvature of the function mapping ΔP onto the mean of the subjective contingency distribution. For reference, if the data from Tables 2 and 4 are grouped together, the average value of the

quadratic term is 0.29, 95% CI [0.22, 0.37].

The fifth conclusion is where discussing both experiments together is the most informative as it concerns the variation in the standard deviation of the subjective contingency distribution as a function of ΔP. Taken apart, each experiment provides only weak support for the existence of a quadratic trend in the relation between ΔP and the standard deviation. That claim becomes much stronger once the two experiments are put together as Experiment 2 replicates the finding of Experiment 1. The quadratic term for the polynomial regression once Tables 3 and 5 have been combined is 0.10, 95% CI [0.04, 0.15]. Figure 8 shows the standard deviation of the subjective mean distribution as a function of (recoded) ΔP as well as the best-fitting polynomial equation ($y = 0.10x^2 - 0.01x + 1.00, r^2 = 0.96$) once the data from Figures 4 and 7 have been combined. The conclusion in both cases is the same: though weak in amplitude, the quadratic trend is clearly present. On the other hand, it seems unlikely that there is a linear trend in the relation between ΔP and the standard deviation. The average linear term for the polynomial regression once Tables 3 and 5 have been combined is -0.01, 95% CI [-0.05, 0.04], which can be compared to the linear term of -0.01 shown in the polynomial equation fitted to the data in Figure 10.

General discussion

In this study, participants were exposed to flow of stimuli (i.e., a target cue and an outcome) and had to judge whether an outcome was more (Experiment 1) or less (Experiment 2) likely to appear following a cue. Our goal was to test three hypotheses one could make when applying SDT to contingency assessment: (a) the subjective contingency is drawn from a Gaussian distribution; (b) the mean of the distribution is a linear function of ΔP; (c) its standard deviation is constant.

We found that the data could be accounted for by a SDT Gaussian model. On the other hand, the assumption that the mean of the subjective contingency distribution is linearly

related to ΔP does not seem to hold. In both experiments, the participants were more sensitive to positive contingencies than to negative ones. Finally, the assumption that the standard deviation of the subjective contingency distribution remains constant over the ΔP value also needs qualification. The perception of the cue-outcome contingency becomes more variable as the contingency becomes more negative or positive, though only slightly. For practical purposes, if the SDT model is used only to measure how sensitive a participant is to variations in ΔP, the quadratic trend in the relation between ΔP and the standard deviation of the subjective contingency distribution can probably be ignored, especially if the comparison does not involve extremely positive or negative contingencies. However, any theoretical account of the variability of the subjective contingency must take it into account since, albeit small, it is no less real.

We reached these conclusions using the streamed-trial procedure, a new experimental paradigm created by Allan et al. specifically to allow the application of SDT to associative learning. The version Allan et al. used (with the notable exception of Experiment 1b in Allan et al, 2008) involved a simultaneous conditioning procedure, with the cue and outcome both shown at the same time, while we used a delay conditioning paradigm with the outcome following the cue. Despite this difference, we were able to replicate the effect of ΔP demonstrated by Allan et al. This strengthens the generality of the streamed-trial procedure as a tool for studying human contingency learning.

Though the categorical dependent variable we used in this study might seem inappropriate in regard to the quantitative goal we set ourselves, the results show that rich quantitative information can be derived from it once it is analyzed through SDT. Nevertheless, one may wonder whether the same conclusion would have been reached or not using more traditional techniques. For instance, we could have used a 6-point Likert scale ranging from -3 to +3 or have recorded contingency judgement by having participants pick a

location on a continuous line ranging from -100 to +100.

Our dependent measure is equivalent in fact to a 6-point Likert scale, except that it is labeled from "No – Very Sure" to "Yes – Very Sure" rather than from -3 to +3. No matter how it is labeled, it remains an ordinal measure of performance and SDT is one appropriate way to analyze it. This only requires commitment to the assumptions underlying SDT, i.e. performance is based on an internal continuous decision variable whose comparison to decision thresholds determines the response, both of which are explicit, minimal and quite plausible when applied to situations like the one studied here.

By contrast, the standard practice of analyzing Likert scale data as if they were numerical variables relies on a set of implicit and quite improbable assumptions. It is bound to introduce distortions in the conclusions which make it inappropriate for quantitative investigations. As a case in point, we reanalyzed the data from Experiment 1 as if we had used a Likert scale ranging from -3 to +3. By treating these data as if they were numerical variables, we were able to compute the equivalent of the psychometric function in Figure 3, showing how the average Likert scale rating varies with the recoded ΔP (Figure 9). This function (well-fitted by $y = 1.20x^2 + 1.06x - 0.60, r^2 = 0.94$) over-estimates the quadratic trend and would have led to us to conclude wrongly that the participants were *insensitive* to the negative contingencies.

Using a continuous scale to assess contingency learning has the advantage over a Likert scale that the dependent variable is genuinely numerical. However, it comes with its own set of implicit assumptions which make it inappropriate for the kind of quantitative investigations we conducted in the present study. Let I be the objective stimulus intensity, X the subjective stimulus intensity and Y the point on the continuous line selected by the participant. X and I are related by a function F so that X = F(I). Y and X are related by a function S so that Y = S(X) = S[F(I)] = R(I). The goal of this study was to reveal the

properties of F. Unfortunately, from the observation of the Y = R(I) relation, it is impossible

to infer valuable information about function F without making assumptions about function S.

There are no theoretical guidelines to constrain these assumptions which reflect the fact that

the very process by which the participants were able to map the subjective intensity of the

stimulus onto a point on the continuous scale remains mysterious and was not properly

investigated.

In conclusion, we consider that neither a Likert scale treated as a numerical variable

nor the use of a continuous scale provides a substitute for SDT. Only the latter is grounded in

theoretical principles sound enough to guarantee the validity of the quantitative conclusions

derived from it.

A surprising outcome of the present investigation is the discovery that the participants

were less sensitive to negative contingencies than to positive ones. There are some

corroborations for this finding in the literature. In an operant task, Wasserman, Elek,

Chatlosh & Baker (1993) reported that the rating of the contingency relation between the cue

(a key-tapping response) and the outcome (a flash of light) was more sensitive to the

probability of the outcome following the cue than of the outcome in the absence of the cue.

Data from Mutter & Williams (2004) collected in another operant procedure shows that the

participants, especially the older ones, were less sensitive to change in the negative

contingencies than in the positive ones. Finally, using a streamed-trial procedure, Allan et al.

(2008) asked participants to judge how strong the relation between a cue and an outcome

was. They found that, for a given absolute value of ΔP, participants judged that the relation

between the cue and the outcome was stronger when the contingency was positive rather than

when it was negative. They interpreted this result as reflecting an effect on the decision

criterion λ rather than on the sensitivity to the contingencies. The present study shows that

this explanation is not sufficient: the psychometric curves in Figures 3 and 5 are free of any

influence from the decision criterion yet still display an asymmetry between positive and negative contingencies.

Moreover, the analysis conducted by Allan et al. (2008) was based on the fit of a Gaussian cumulative function to psychometric curves similar to the one displayed in Figure 1, whose parameters are then analyzed within the framework of a Gaussian SDT model with constant variance. This is not enough to distinguish between an account assuming a change in the decision criterion but not in the underlying subjective contingency distributions, one account assuming the reverse and the other assuming both. Suppose, for instance, that in a given condition, the probability for the participant to detect a contingency between the cue and the outcome is equal to $p1$. In that condition, the subjective contingency is drawn from a Gaussian distribution with mean μ_1 and standard deviation σ. If the subjective contingency falls above a criterion λ_1, the participant perceives a positive contingency. The stimulus condition is changed and the probability to detect a positive contingency falls to $p2$. This could be explained either by assuming that the criterion moved from λ_1 to λ_2 (with $\lambda_2 - \lambda_1 = k$) so that

$$
\begin{aligned}
p2 \quad &= \quad \phi\left(-\frac{\lambda_2 - \mu_1}{\sigma}\right) \\
&= \quad \phi\left(-\frac{\lambda_1 + k - \mu_1}{\sigma}\right)
\end{aligned}
$$

or that μ_1 has moved to μ_2 ($\mu_1 - \mu_2 = k$) so that

$$
\begin{aligned}
p2 \quad &= \quad \phi\left(-\frac{\lambda_2 - \mu_1}{\sigma}\right) \\
&= \quad \phi\left(-\frac{\lambda_1 + k - \mu_1}{\sigma}\right)
\end{aligned}
$$

Between these two extremes, we can build an infinity of mathematically equivalent models assuming a change in both the decision criterion and the mean of the distribution.

The same problem undermines the conclusion of Siegel et al. (2009) that cue competition effects in the streamed-trial procedure reflect a performance deficit, i.e. a change

in the decision criterion, rather than a learning deficit, i.e. a change in the subjective contingency distribution. A study by Laux et al. (2010) seems to provide stronger evidence for this conclusion. Using a streamed-trial procedure, they manipulated the contingency between a target cue and the outcome, which could be null or positive, and between a competing cue and the outcome, which could also be null (weak alternative condition) or positive (strong alternative condition). Participants were more likely to detect the positive contingency between the target cue and the outcome in the weak alternative condition than in the strong alternative condition. Laux et al. (2010) applied a Gaussian SDT model with constant variance to their data, making the usual assumptions. In the null (respectively positive) contingency condition, the subjective contingency was drawn from a Gaussian distribution with mean μ_1 (respectively μ_2) and standard deviation σ; if the subjective contingency was above a threshold λ, the subject perceived a contingency between the target cue and the outcome, otherwise he did not. Laux et al. (2010) showed that d', the distance between μ_1 and μ_2, was roughly the same in the weak alternative and in the strong alternative condition, while λ was much higher in the latter condition than in the former one. However, as all the parameters in an SDT model are expressed relative to a reference condition (in this case, μ_1) an alternative explanation, assuming that μ_1 and μ_2 changed while λ remained constant, is possible as long as the distance between μ_1 and μ_2 remains the same. This latter requirement is not as far-fetched as it might seem as simulations of the procedure of Laux et al. using the Rescorla-Wagner simulator developed by Mondragon, Alonso, Fernández, & Gray (2013) show that this is exactly what the Rescorla-Wagner model predicts. The associative strength of a contingent and of a non-contingent cue are lower in the strong alternative condition than in the weak alternative one but the distance between the two associative strengths remains the same.

On the other hand, the claim by Allan et al. (2005) that the outcome density affects

the decision criterion λ and not the underlying subjective contingency distribution as well as the conclusion by Perales et al. (2005) that this is also the case for the cue frequency are not susceptible to the arguments we used against Siegel et al. (2009) and Laux et al. (2010). As such, they seem to be based on more solid grounds.

Overall, this paper complements previous work by Allan et al. in showing the value of SDT for contingency learning, and beyond that, of a psychophysical approach to associative learning. Psychophysics is concerned with how objective stimulus intensity maps onto subjective stimulus intensity. If we consider that the objective contingency between two stimuli can be measured objectively with ΔP, asking how the objective contingency maps onto the subjective one is an obvious question if associative learning is approached from a psychophysical perspective. This issue has, on the other hand, attracted very little attention in traditional modeling of associative learning. Another issue central to psychophysics is the one of noise. An important aim of most psychophysical models is to characterize the noise in the stimulus representation. By contrast, most models of associative learning are purely deterministic and have never considered possible sources of noise in associative learning nor the impact it could have on performance.

By tackling directly the issue of the mapping of objective contingency onto subjective contingency and the question of noise in subjective contingency, studies such as this one pave the way for a new type of model of associative learning rooted in psychophysics. Nevertheless, further research is needed to bring more empirical grist to the theoretical mill of such yet-to-be models. For instance, at this point, we have no explanation for the higher sensitivity to positive contingencies or the increase in the variability of the perceived cue-outcome contingency as the contingency becomes more positive or negative. This requires further investigation.

Intuitively, the higher sensitivity to positive contingencies could be explained by

assuming that the participants were more sensitive to things happening (in this case, the outcome being presented following the cue) than to things not happening (the outcome not being presented following the cue). In support of this hypothesis, one could point out that the error-correction signal identified in the midbrain dopamine neurons is stronger and more graded when an unexpected reward occurs than when an expected reward is omitted (for instance, see Fiorillo, Tobler, & Schuttz, 2003 or Stauffer, Lak, & Schultz, 2014). But it is very unlikely that the dopamine system is involved in the present study given that the stimuli used in the procedure lacked any rewarding properties. Moreover, the idea that the participants are differentially sensitive to the occurrence or non-occurrence of an event runs against Gallistel and Gibbon's (2000) claim that, at least in pigeon autoshaping, extinction proceeds at the same speed, if not faster, than acquisition. At this point, the only conclusion that can be reached without any doubt is that the higher sensitivity of the participants to the positive contingency does not fit with the Rescorla-Wagner model which predicts a linear relation between ΔP and the subjective contingency (Rescorla & Wagner, 1972; Danks, 2003).

Many other questions remain unanswered. For instance, ΔP has two components: the probability of the outcome in the presence (or just following) the cue, which corresponds to the number of pairings between the cue and the outcome, and the probability of the outcome in the absence of the cue, i.e. the number of unpaired presentations of the outcome. In the present study, both probabilities changed when we manipulated ΔP in order to keep the outcome density constant between streams for reasons already discussed above. However, that means that we cannot know how much either of the components of ΔP (number of cue-outcome pairing and unpaired presentation of the outcome) contributed to the subjective contingency. Did they both contribute and if so, in which proportion? Or did only one contribute and if so, which one? This is certainly a question worth investigating.

The considerable individual differences observed in the data is another issue that deserves further study. The linear term of the polynomial equation shown in Tables 2 and 3 captures the sensitivity of participants to a change in the cue-outcome contingency. As can be seen, it varied greatly among them. We already highlighted the individual differences shown in Tables 3 and 5 in the pattern of variation of the standard deviation as a function of ΔP. Future research could investigate whether these differences can be predicted based on some characteristics of participants and whether they are themselves predictors of their performance in specific situations. For instance, one would expect that participants with a higher sensitivity to the contingencies would be more responsive to extinction treatment than those with a lower sensitivity to the contingencies.

Finally, it would be fundamental to study the impact of several variables critical for associative learning in the SDT model, such as the amount of training, or temporal variables (duration of stimuli, time interval between them and between trials, etc.). Would these variables have an impact on the mean of the distribution, on their standard deviation, or on both? Furthermore, it would be interesting to know whether, keeping ΔP and its component probabilities constant, the performance of the participants varies with the number of trials in a stream, whether it changes with further exposure to the task or whether sensitivity to the contingencies can be improved by providing participants with feedback about their performance. We mention all these issues not merely to point out the obvious limitations of the current study but to show the exciting new pathways that can be trodden within a psychophysical approach to associative learning and which could lead us to a deeper and more fundamental understanding of associative learning in both men and animals.

References

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

Allan, L. G., Hannah, S. D., Crump, M. J. C., & Siegel, S. (2008). The psychophysics of contingency assessment. *Journal of Experimental Psychology: General, 137,* 226-243.

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

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

Danks, D. (2003). Equilibria of the Rescorla-Wagner model. *Journal of Mathematical Psychology, 47,* 109-121.

Fiorillo, C. D., Tobler, P. N., & Schultz, W. (2003). Discrete coding of reward probability and uncertainty in dopamine neurons. *Science, 299,* 1898-1902.

Gallistel, C. R., & Gibbon, J. (2000). Time, rate, and conditioning. *Psychological Review, 107,* 289-344.

Hannah, S. D., Crump, M. J. C., Allan, L. G., & Siegel, S. (2009). Cue-interaction effects in contingency judgments using the streamed-trial procedure. *Canadian Journal of Experimental Psychology, 63,* 103-112.

Hannah, S. D., & Allan, L. G. (2011). The criterion-calibration model of cue interaction in contingency judgments. *Learning & Behavior, 39,* 171-190.

Harvey, L. O. (2013). *Parameter estimation of signal detection models: RscorePlus user's manual* (Tech. Rep.). Boulder, Colorado 80309-0345 USA http://psych.colorado.edu/~lharvey/Software%20Zip%20Files/RscorePlus_doc_5_8_3.pdf. Department of Psychology and Neuroscience University of Colorado.

Laux, J.P., Goedert, K.M., & Markman, A.B. (2010). Causal discounting in the

presence of a stronger cue is due to bias. *Psychonomic Bulletin & Review, 17,* 213-218.

Macmillian, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide.* Cambridge: Cambridge University Press.

Mackintosh , N. J. (1975). A theory of attention: Variations in the associability of stimuli with reinforcement. *Psychological Review, 82,* 276-278.

Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *Journal of the Society for Industrial & Applied Mathematics, 11,* 431-441.

Mondragon, E., Alonso, E., Fernandez, A., & Gray, J. (2013). An extension of the Rescorla and Wagner simulator for context conditioning. *Computer Methods and Programs in Biomedecine, 110,* 226-230.

Mutter, S.M., & Williams, T.W. (2004). Aging and the detection of contingency in causal learning. *Psychology and Aging, 19,* 13-26.

Miller, R. R. & Matzel, L. D. (1988). The comparator hypothesis: A response rule for the expression of associations. In G. H. Bower (Ed), *The Psychology of Learning and Motivation* (Vol. 22, pp. 51-92). San Diego, CA: Academic Press.

Pagès, J. (2010). *Statistique générale pour utilisateurs, vol I: Méthodologie.* Rennes, France: Presse Universitaire de Grenoble.

Pearce, J. M., & Hall, G. (1980). A model for Pavlovian learning: Variations in the effectiveness of conditioned but not unconditioned stimuli. *Psychological Review, 87,* 532-552.

Perales, J. C., Catena, A., Shanks, D. R., & González, J. A. (2005). Dissociation between judgments and outcome-expectancy measures in covariation learning: A signal detection theory approach. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31,* 1105-1120.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2002). *Numerical recipes in C++: The art of scientific computing (2nd edition)*. Cambridge: Cambridge University Press.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., & Flannery, B. P. (2007). *Numerical recipes: The art of scientific computing (3rd edition)*. New York: Cambridge University Press.

Rescorla, R. A. (1967). Pavlovian conditioning and its proper control procedures. *Psychological Review, 74,* 71-80.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & F. Prokasy (Eds), *Classical Conditioning II* (pp. 64-99). New York: Appleton Century Crofts.

Schanks, D. R. (2007). Associationism and cognition: Human contingency learning at 25. *Quarterly Journal of Experimental Psychology, 60,* 291-309.

Siegel, S., Allan, L. G., & Hannah, S. D. (2009). Applying signal detection theory to contingency assessment. *Comparative Cognition & Behavior Reviews, 4,* 116-134.

Stauffer, W. R., Lak, A., & Schultz, W. (2014). Dopamine reward prediction error reflects marginal utility. *Current Biology, 24,* 2491-2500.

Vadillo, M. A., Miller, R. R., & Matute, H. (2005). Causal and predictive-value judgments, but not predictions, are based on cue-outcome contingency. *Learning & Behavior, 33,* 172-183.

Wagner, A. R. (1981). SOP: A model of automatic memory processing in animal behavior. In N. E. Spear & R. R. Miller (Eds), *Information processing in animals: Memory mechanisms* (pp. 5.47). Hillsdale, NJ: Erlbaum.

Wasserman, E.A., Elek, S.N., Chatlosh, D.L., & Baker, A.G. (1993). Rating causal relations: Role of probability in judgments of response-outcome contingency. *Journal of Experimental Psychology: Learning, Memory and Cognition, 19,* 174-188.

Wickens, T. D. (2002). *Elementary signal detection theory.* New York: Oxford University Press.

Appendix A

Average proportion of responses per response category (yes/no answer + confidence level) as a function of ΔP in Experiment 1. Numbers in parenthesis are the uncorrected standard deviations.

	NO			YES		
	--			--		
ΔP/category	Very Sure	Sure	Not Sure	Not Sure	Sure	Not Sure
-0.8	0.17 (0.22)	0.34 (0.19)	0.25 (0.20)	0.10 (0.08)	0.12 (0.12)	0.02 (0.05)
-0.4	0.09 (0.13)	0.35 (0.20)	0.27 (0.19)	0.13 (0.09)	0.14 (0.12)	0.02 (0.05)
0	0.05 (0.06)	0.25 (0.14)	0.28 (0.17)	0.21 (0.10)	0.19 (0.13)	0.03 (0.05)
0.4	0.02 (0.03)	0.13 (0.07)	0.18 (0.08)	0.29 (0.16)	0.32 (0.11)	0.07 (0.08)
0.8	0.00 (0.01)	0.05 (0.05)	0.09 (0.06)	0.25 (0.20)	0.40 (0.16)	0.21 (0.22)

Appendix B

Average proportion of responses per response category (yes/no answer + confidence level)

as a function of ΔP in Experiment 2. Numbers in parenthesis are the uncorrected standard

deviations.

	YES			NO		
	--			--		
ΔP/category	Very Sure	Sure	Not Sure	Not Sure	Sure	Not Sure
-0.8	0.15 (0.20)	0.35 (0.18)	0.17 (0.13)	0.15 (0.09)	0.16 (0.12)	0.02 (0.02)
-0.4	0.10 (0.20)	0.30 (0.18)	0.23 (0.13)	0.19 (0.09)	0.17 (0.12)	0.01 (0.05)
0	0.04 (0.12)	0.25 (0.15)	0.23 (0.14)	0.23 (0.11)	0.21 (0.11)	0.03 (0.02)
0.4	0.02 (0.03)	0.13 (0.10)	0.17 (0.09)	0.26 (0.16)	0.33 (0.15)	0.09 (0.18)
0.8	0.00 (0.01)	0.08 (0.10)	0.10 (0.08)	0.20 (0.19)	0.39 (0.20)	0.22 (0.28)

Table 1

Number of trials as a function of trial type and ΔP condition along with resulting

probabilities

ΔP	AX+	AX-	A+	A-	P(O|AX)	P(O|A)	P(O)	P(X)
-0.8	1	9	9	1	0.10	0.90	0.5	0.5
-0.4	3	7	7	3	0.30	0.70	0.5	0.5
0	5	5	5	5	0.50	0.50	0.5	0.5
0.4	7	3	3	7	0.70	0.30	0.5	0.5
0.8	9	1	1	9	0.90	0.10	0.5	0.5

Note - AX+ cue and outcome (O) are presented. AX- cue is presented but not outcome. A+ outcome is presented but not cue. A- Neither cue nor outcome are presented.

Table 2

Parameters of second-degree polynomial $ax^2 + bx + c$ that, for each participant in

Experiment 1, was fitted to function mapping (recoded) ΔP value over mean of subjective

contingency distribution.

Participants/Parameters	a	b	c	r^2
P1	0.44	1.47	0.01	1.00
P2	0.17	0.54	- 0.05	0.95
P3	0.52	2.32	- 0.10	0.99
P4	0.30	0.45	0.03	0.99
P5	0.41	1.86	-0.06	1.00
P6	0.14	0.42	-0.01	1.00
P7	0.14	0.21	0.13	0.65
P9	0.28	1.49	0.02	0.99
P10	0.15	0.57	-0.03	1.00
P11	-0.06	0.60	0.07	0.94
P12	0.41	1.11	0.02	1.00
P13	0.15	0.39	0.08	0.97
P15	0.61	0.99	- 0.07	0.98
P16	0.59	0.97	0.14	0.95
P17	0.19	1.49	0.03	0.99
Mean	0.30	0.99	0.01	0.99
95% CI	[0.19, 0.40]	[0.65, 1.34]	[-0.03, 0.05]	[0.97, 1.00]

Table 3

Parameters of second-degree polynomial $ax^2 + bx + c$ that, for each participant in

Experiment 1, was fitted to function mapping (recoded) ΔP value over standard deviation of

subjective contingency distribution.

Participants/Parameters	a	b	c	r^2
P1	0.02	0.11	0.95	0.74
P2	0.20	0.04	0.89	0.57
P3	0.43	-0.02	1.04	0.74
P4	-0.06	-0.08	0.94	0.31
P5	0.41	0.28	1.06	0.97
P6	-0.03	0.00	0.97	0.05
P7	0.05	-0.23	0.99	0.98
P9	0.02	0.00	1.03	0.04
P10	-0.07	-0.08	1.06	0.74
P11	-0.02	-0.07	1.06	0.28
P12	0.04	0.26	1.15	0.79
P13	0.02	-0.14	0.92	0.77
P15	0.26	0.17	0.87	0.64
P16	0.39	0.07	1.04	0.78
P17	0.17	0.08	1.03	0.65
Mean	0.12	0.03	1.00	0.72
95% CI	[0.02, 0.22]	[-0.05, 0.10]	[0.96, 1.04]	[0.51, 0.85]

Table 4

Parameters of second-degree polynomial $ax^2 + bx + c$ that, for each participant in

Experiment 2, was fitted to function mapping (recoded) ΔP value over mean of subjective

contingency distribution.

Participants/Parameters	a	b	c	r^2
P1	-0.02	1.71	0.05	1.00
P3	0.33	1.85	0.09	1.00
P4	0.33	0.34	- 0.06	0.98
P5	0.39	0.25	-0.07	0.99
P6	0.08	0.59	0.04	0.95
P7	0.67	1.07	-0.01	0.99
P8	0.30	0.38	0.21	0.83
P9	0.16	0.34	-0.13	0.90
P10	0.14	0.77	0.21	0.93
P11	0.46	0.83	0.08	0.99
P12	0.58	0.74	0.06	0.97
P13	0.38	0.37	0.03	0.87
P14	0.07	0.14	0.02	0.79
P15	0.61	1.69	-0.10	0.99
P16	0.03	0.06	0.06	0.46
P17	0.31	0.64	0.13	0.97
Mean	0.30	0.74	0.04	0.97
95% CI	[0.19, 0.41]	[0.43, 1.04]	[-0.01, 0.09]	[0.93, 0.99]

Table 5

Parameters of second-degree polynomial $ax^2 + bx + c$ that, for each participant in

Experiment 2, was fitted to function mapping (recoded) ΔP value over standard deviation of

subjective contingency distribution.

| Participants/Parameters	a	b	c	r^2
P1	0.17	-0.03	1.04	0.87
P3	-0.06	-0.07	1.36	0.03
P4	-0.04	-0.01	1.01	0.07
P5	0.09	-0.01	0.96	0.14
P6	0.18	-0.17	0.97	0.71
P7	0.11	0.20	0.94	0.84
P8	-0.03	-0.01	0.93	0.07
P9	0.14	-0.06	1.05	0.75
P10	0.12	-0.02	0.96	0.21
P11	0.12	-0.12	1.10	0.35
P12	0.04	-0.01	0.93	0.05
P13	-0.13	-0.09	0.93	0.76
P14	0.04	-0.06	0.95	0.63
P15	0.06	0.00	1.01	0.07
P16	0.35	-0.02	0.98	0.98
P17	0.11	-0.14	0.90	0.69
Mean	0.08	-0.04	1.00	0.56
95% CI	[0.02, 0.14]	[-0.08, 0.01]	[0.94, 1.06]	[0.30, 0.75]
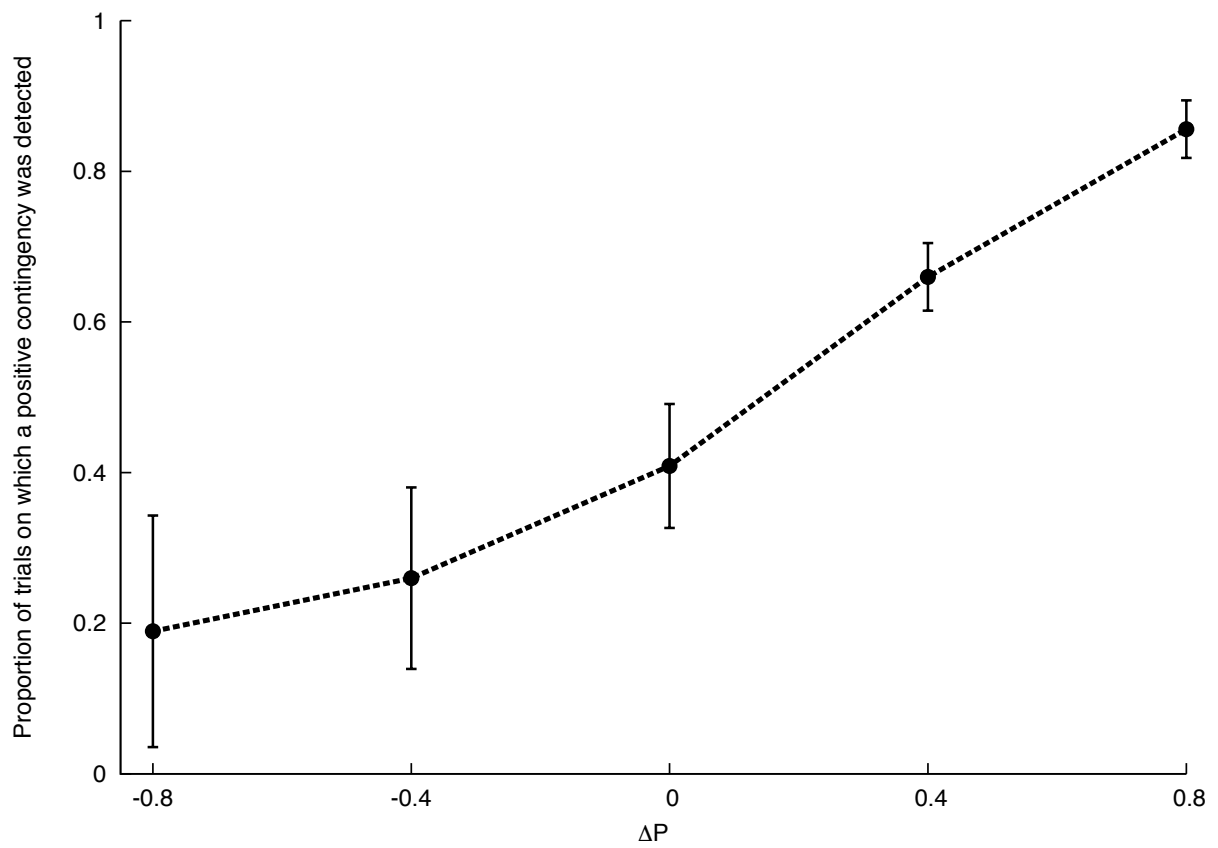
Figure 1. Average proportion of trials on which positive contingency was detected as a function of ΔP in Experiments 1. Error bars are 95% CI.
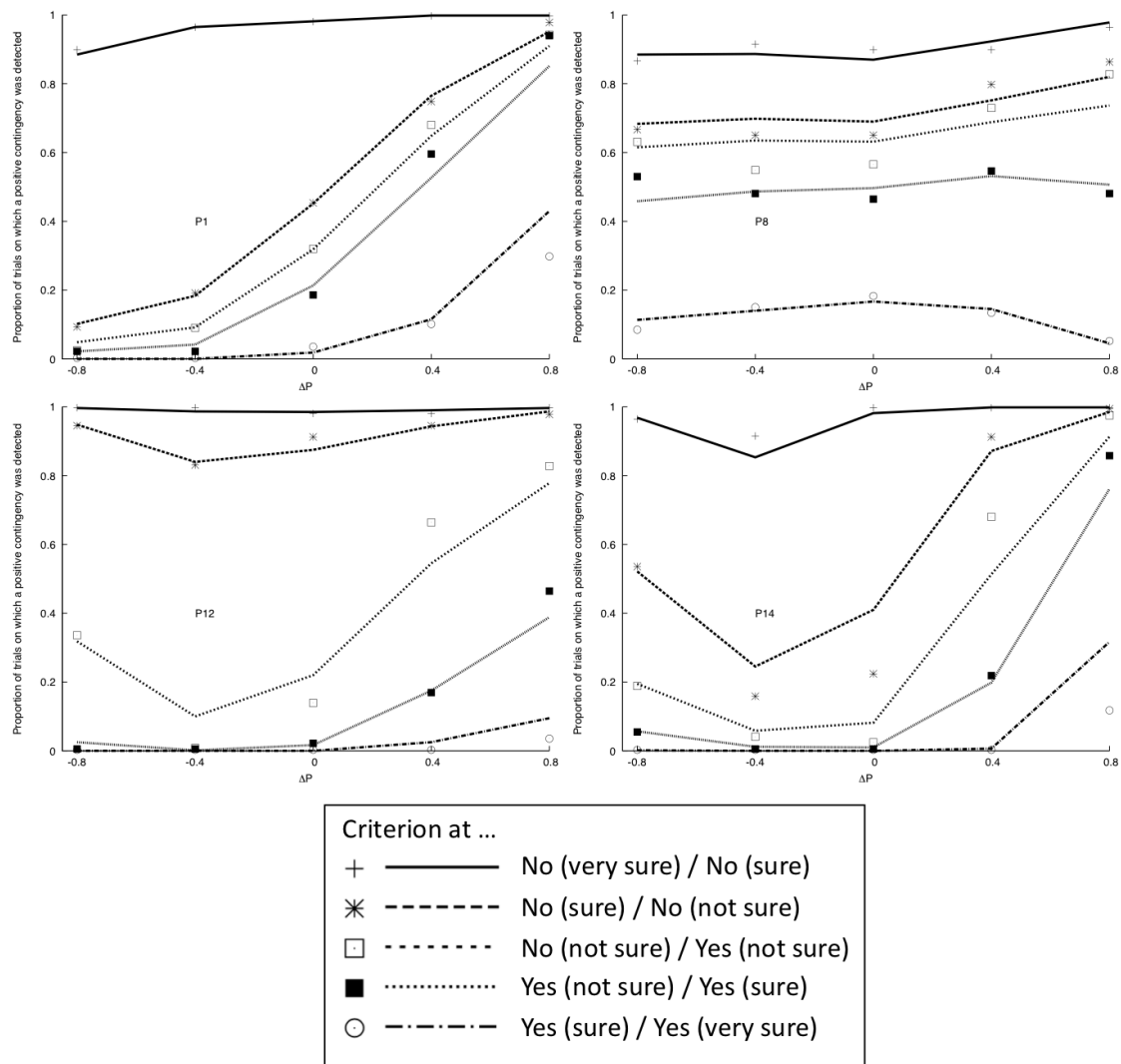
Figure 2. Proportion of trials, as a function of ΔP, on which a positive contingency would

have been detected had the decision threshold been located at one of the various confidence

thresholds for four participants in Experiment 1. Points are actual data while curves are

predictions of SDT model. A χ^2 adjustment test indicated that performance of these

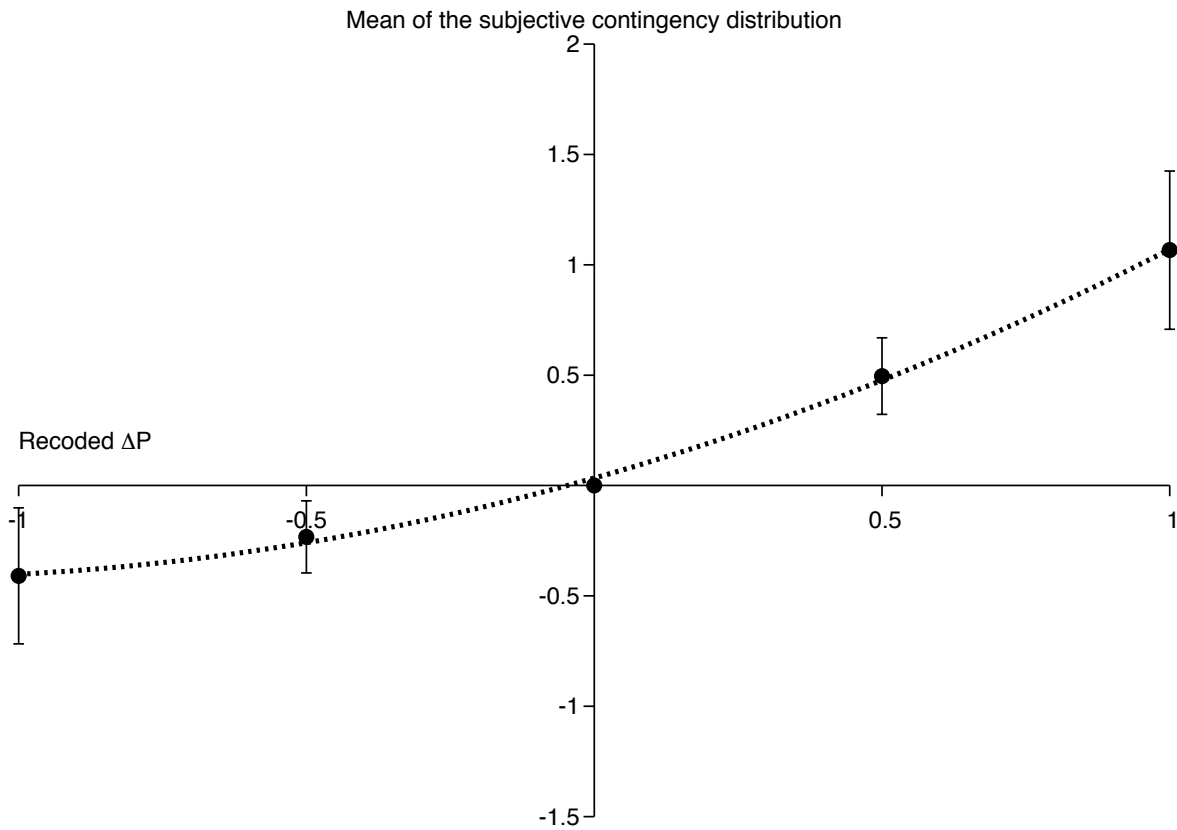participants deviated significantly from SDT model.
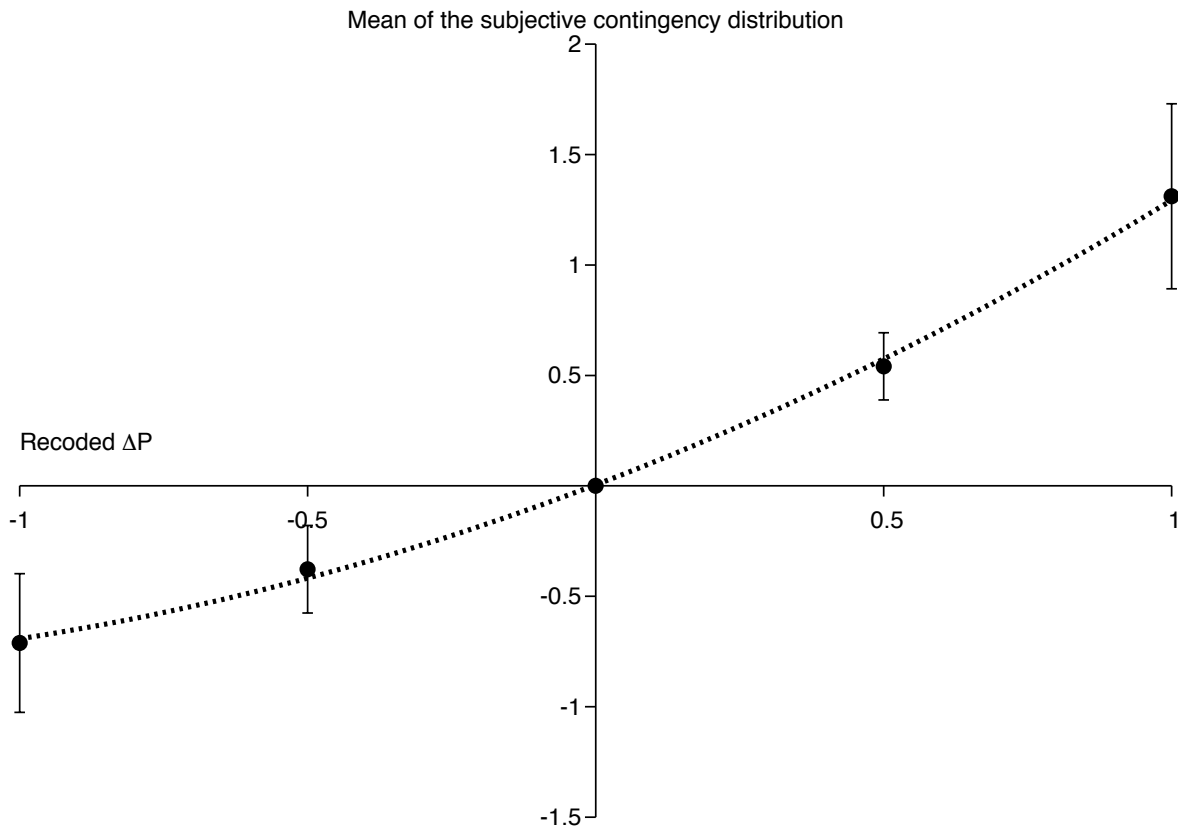
Figure 3. Average mean of subjective contingency distribution as a function of (recoded) ΔP in Experiment 1. Error bars are 95% CI. The best-fitting second-order polynomial ($y = 0.30x^2 + 0.99x + 0.01, r^2 = 1.00$) is also shown.
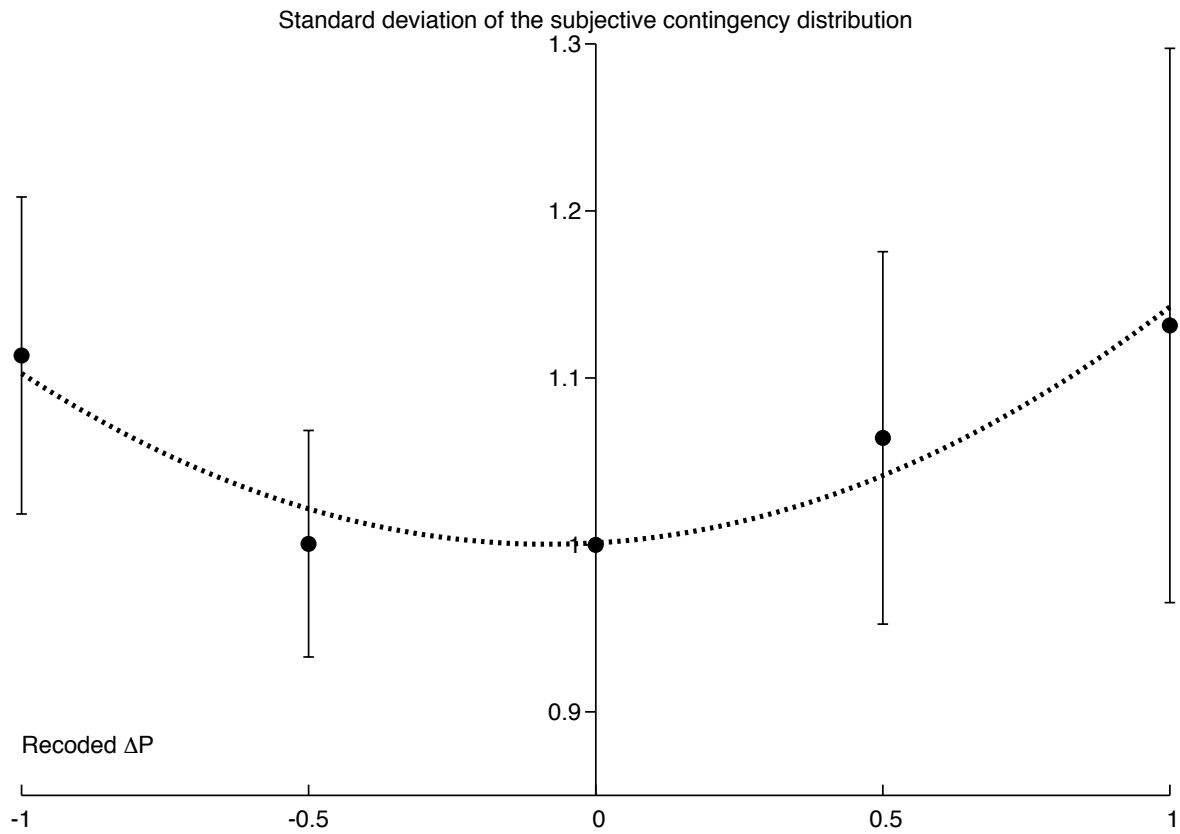
Figure 4. Average standard deviation of subjective contingency distribution as a function of

(recoded) ΔP in Experiment 1. Error bars are 95% CI. The best-fitting second-order

polynomial ($y = 0.12x^2 + 0.02x + 1.00, r^2 = 0.92$) is also shown.
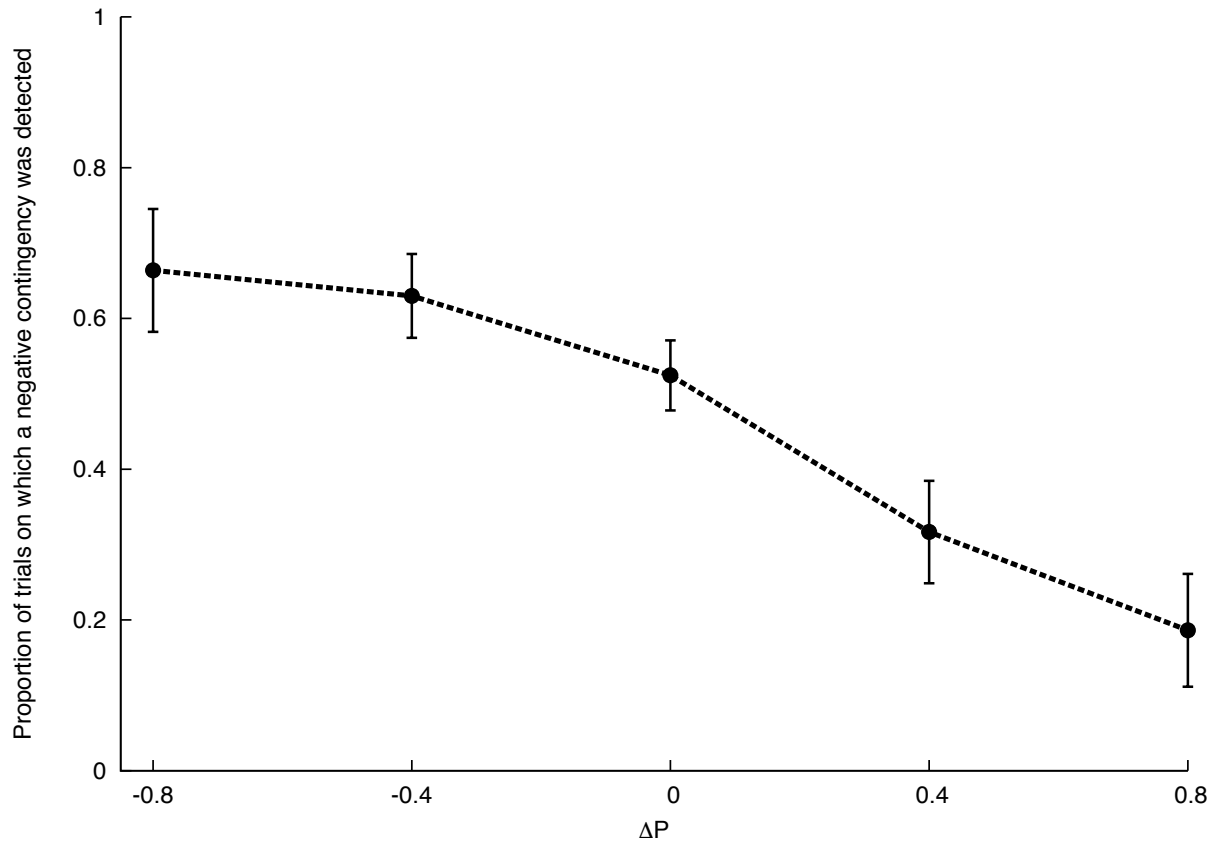
Figure 5. Average proportion of trials on which a negative contingency was detected as a

function of ΔP in Experiments 2. Error bars are 95% CI.

Figure 6. Average mean of subjective contingency distribution as a function of (recoded) ΔP in Experiment 2. Error bars are 95% CI. The best-fitting second-order polynomial ($y = 0.30x^2 + 0.74x + 0.03, r^2 = 1.00$) is also shown.
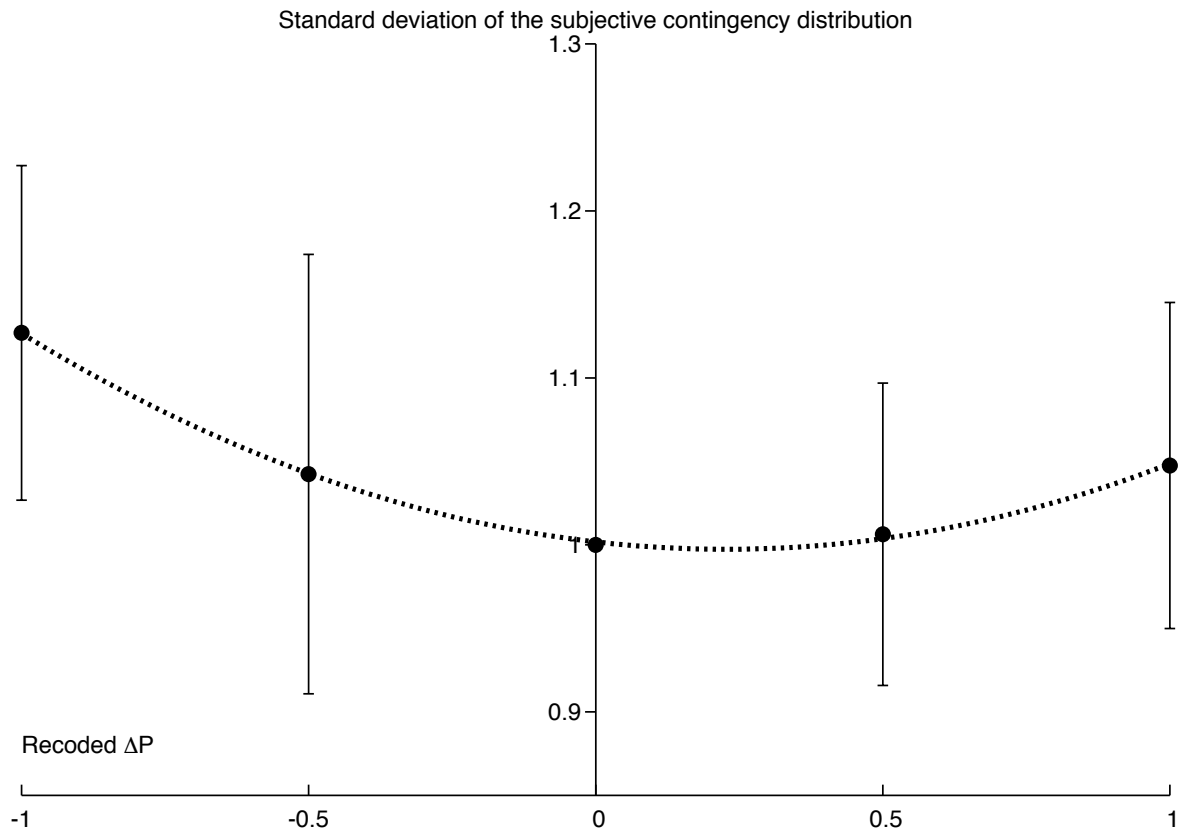
Figure 7. Average standard deviation of subjective contingency distribution as a function of (recoded) ΔP in Experiment 2. Error bars are 95% CI. The best-fitting second-order polynomial ($y = 0.09x^2 - 0.04x + 1.00, r^2 = 1.00$) is also shown.
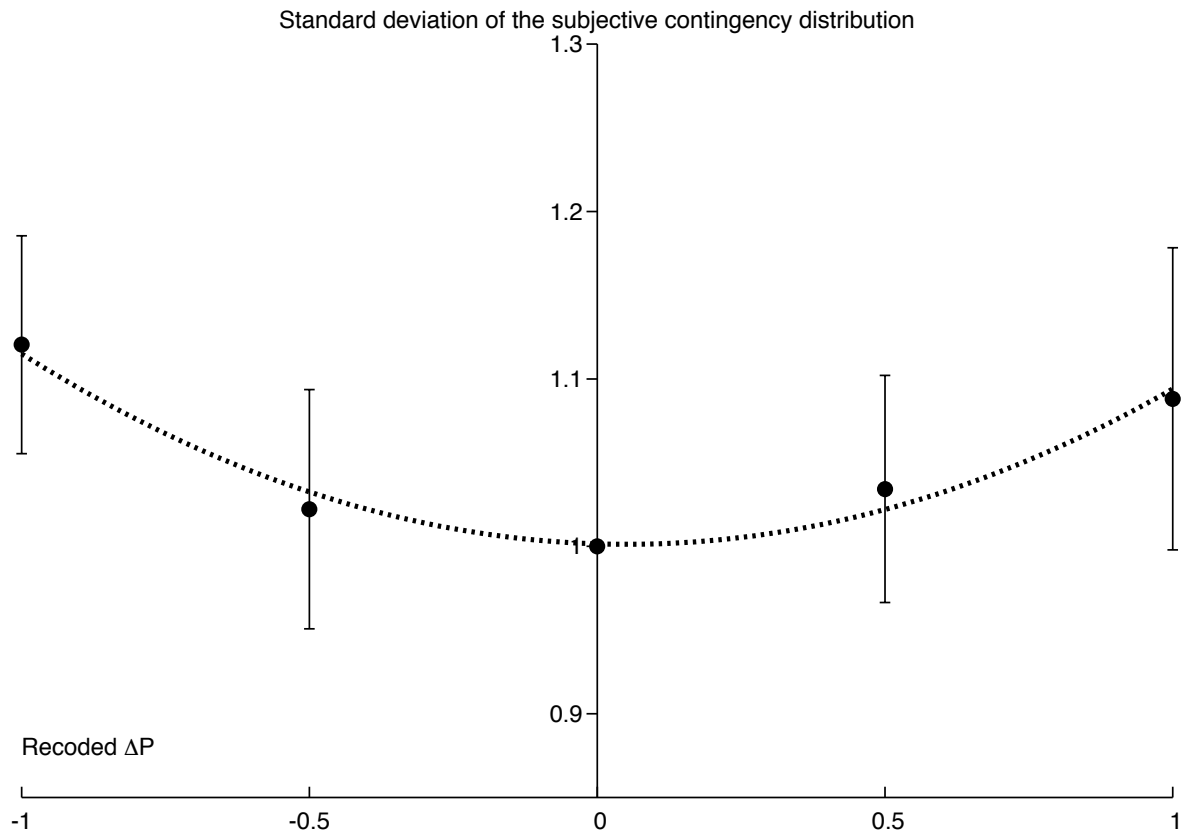
Figure 8. Average standard deviation of subjective contingency distribution as a function of

(recoded) ΔP in Experiments 1 and 2 combined. Error bars are 95% CI. The best-fitting

second-order polynomial ($y = 0.10x^2 - 0.01x + 1.00, r^2 = 0.96$) is also shown.
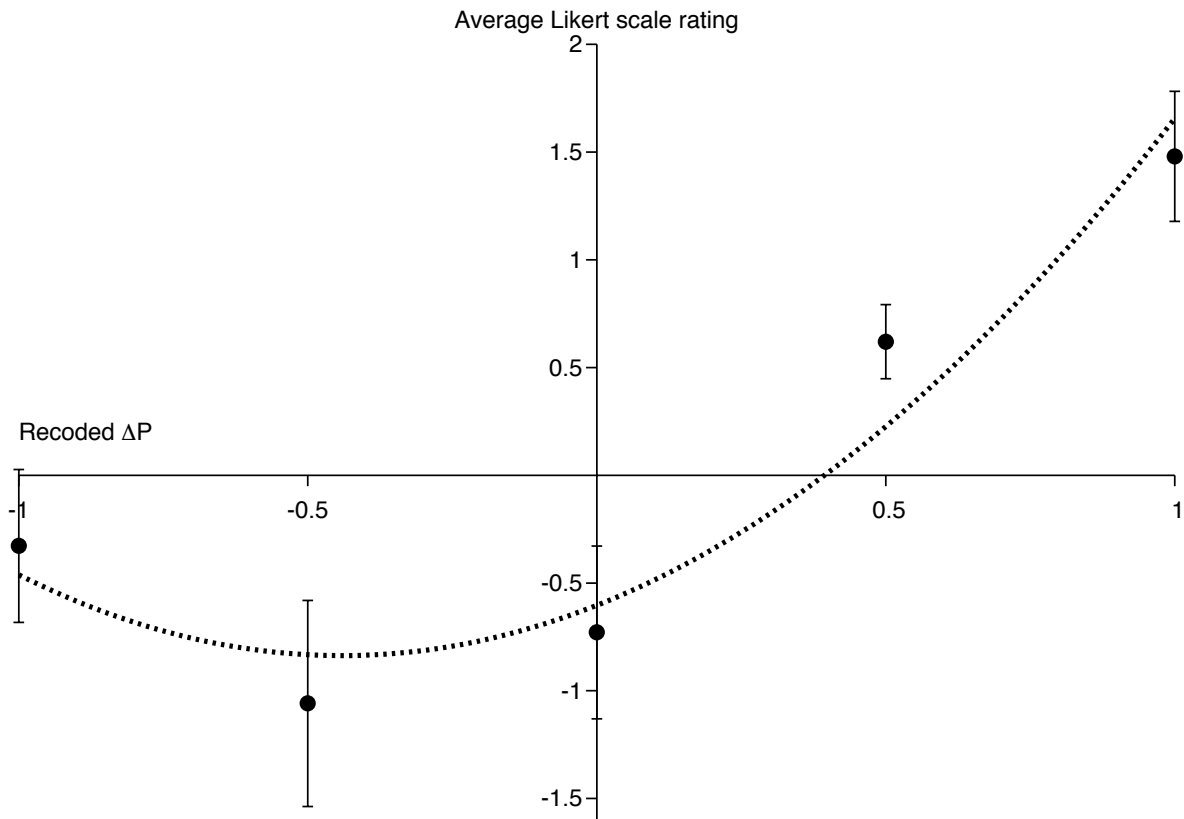
Figure 9. Average Likert scale rating as a function of (recoded) ΔP in Experiment 1. Error

bars are 95% CI. The best-fitting second-order polynomial ($y = 1.20x^2 + 1.06x -$

$0.60, r^2 = 0.94$) is also shown.