

Contingency Judgements

M. J. C. Crump, D. L. G. A. and L. K. H.
McMaster University, Hamilton, Ontario, Canada

The present research was conducted to establish the validity of a novel procedure for measuring human contingency judgements aimed at shortening the length of conventional procedures. Cues and outcomes were simple geometric shapes that were presented in a rapid streaming fashion, reducing the length of a block of trials from several minutes to a few seconds. We establish the reliability of the procedure by replicating two central findings in the contingency judgement literature, and we elaborate on the importance of this method for future research.

The ability to detect and interpret contingencies between events in the environment may be one of the most fundamental processes underlying human cognition and animal learning. Detection of contingencies between events is assumed to mediate the hallmark ability of humans and animals to predict and control events in the environment and, more specifically, is assumed to mediate learning (Gallistel, 2002; Rescorla & Wagner, 1972), categorization and concept formation (Gluck & Bower, 1988), acquisition of causal structure (De Houwer & Beckers, 2002), and decision making (Mandel & Lehman, 1998). Not surprisingly, research into the processes mediating sensitivity to contingency has continued to be a topic of considerable interest for psychologists (for reviews see Allan, 1993; De Houwer &

Beckers, 2002). Given the necessity of processing contingent information across a wide range of domains, it is important to understand the processes contributing to the detection and interpretation of contingent information and the general constraints surrounding their detection and interpretation. Several methods have been previously employed to assess people's knowledge of contingency. We discuss some of the limitations of these techniques and report a novel procedure for measuring knowledge of contingency that is aimed at overcoming previous limitations.

Contingency judgement tasks typically involve rating the strength of relationship between

Correspondence should be addressed to Matthew J. C. Crump, Department of Psychology, Neuroscience & Behaviour, McMaster University, 1280 Main Street West, Hamilton, Ontario, L8S 4K1, Canada. E-mail: crumpmj@mcmaster.ca

The preparation of this paper was supported by research grants from the Natural Sciences and Engineering Research Council of Canada to L. G. A. and by a Natural Sciences and Engineering Research Council of Canada Graduate Scholarship to M. J. C. C. We thank Tom Beckers and one anonymous reviewer for their helpful suggestions, and as well we acknowledge the help provided by Jen Beneteau with regard to data collection. Portions of this research were reported in partial fulfilment of L. K. H.'s undergraduate honours thesis.

binary variables that have been paired over several trials. In general, participants are exposed to a series of trials in which a cue is either presented (C) or not presented (\sim C), and an outcome either occurs (O) or does not occur (\sim O). As a result, there are four possible cue–outcome pairings that can be presented with varying frequencies to manipulate the cue–outcome relationship. Table 1 displays a 2×2 contingency table representing the four different cue–outcome pairings.

The letters inside each cell (A, B, C, D) denote the frequency of occurrence of each cue–outcome pair presented over trials. Conventionally, the contingency between the cue–outcome pairs over trials is defined by the ΔP rule (see Allan, 1980):

$$\Delta P = P(O|C) - P(O|\sim C) = \frac{A}{A+B} - \frac{C}{C+D} \quad (1)$$

The most common variant of the contingency judgement task is the discrete-trials procedure (e.g., Allan & Jenkins, 1983; Dickinson, Shanks, & Evenden, 1984; Shanks, 1985, 1986). Here, each trial consists of one presentation of a cue event (C or \sim C), followed by an outcome event (O or \sim O). For example, the cue could be a fertiliser that is present or absent, and the outcome could be plant growth that occurs or does not occur. At the end of a series of trials participants are asked to rate the strength of relationship between the cue and outcome. Discrete-trial procedures can involve passive instructions, where participants merely observe cue–outcome events over trials, or active (operant) instructions, where participants either initiate a response or do not

initiate a response, and then an outcome occurs or does not occur. It is also common for participants to make predictions about the likelihood of the outcome after the presentation of a cue event on each trial. Finally, at the end of a block of trials the participant is asked to make a judgement about the strength of relationship between the cue and the outcome, usually in the form of a single point estimate (e.g., between -100 and $+100$).

Another common contingency judgement task is the free-operant procedure (e.g., Wasserman, Chatlosh, & Neunaber, 1983). In this case, cue–outcome events are not presented in a discrete, predefined trial-by-trial structure. Instead, participants are free to respond at any time during the experimental interval, and the presentation of an outcome is determined by sampling whether a response was made inside a moving temporal window (e.g., 1 s). Last, although less common, contingency judgement tasks have also presented contingency information to participants in summary format (e.g., Kao & Wasserman, 1993).

• • | | a | o

Importantly, central findings in the contingency judgement literature are generally consistent across tasks. For example, one central finding is that ratings of contingency correlate highly with the programmed (nominal) ΔP values used to describe the cue–outcome pairings. At the same time, there are also conditions under which contingency ratings display systematic departures from ΔP values. For example, it is well known that contingency ratings vary as a function of outcome density (Allan & Jenkins, 1983; Alloy & Abramson, 1979). Specifically, regardless of the actual contingency between a cue and an outcome, ratings of contingency tend to be lower when few outcomes are presented (i.e., low outcome density) and higher when many outcomes are presented (i.e., high outcome density).

The outcome density effect could reveal important constraints about how humans learn about contingencies between events. For example, Alloy and Abramson (1979), in an active version of the discrete-trials task, demonstrated that outcome

1. 2×2 matrix for cue–outcome pairings in a contingency task

	O	\sim O
C	A	B
\sim C	C	D

Note: The letters in each cell (A, B, C, D) represent the joint frequency of occurrence of the four cue–outcome combinations. C: a cue is presented. \sim C: a cue is not presented. O: an outcome occurs. \sim O: an outcome does not occur.

density effects were absent for depressed participants. They suggested that mood could mediate sensitivity to contingencies and that depressed individuals may be “sadder but wiser”. On the other hand, Allan, Siegel, and Tangen (2005) recast contingency judgements in terms of signal detection theory (Green & Swets, 1966) and demonstrated that, at least in the passive contingency task, participants’ sensitivity measure (d') to the contingency did not vary with outcome density, but their bias to report a signal did vary with outcome density (see also Perales, Catena, Gonzalez, & Shanks, 2005). This raises the possibility that outcome density effects, and Alloy and Abramson’s depressive realism effect, reflect a response bias on behalf of the participant. If so, this would still leave much unknown about the processes guiding this response bias. We further explore this issue in the present research and elaborate on its importance in the General Discussion.

a

The length of conventional procedures has placed several constraints on research into contingency judgements. Typically, ratings are given at the end of a block of several trials, which may last anywhere between 5 and 20 minutes. As a result, researchers are faced with investing an inordinate amount of time to collect relatively few estimates of performance across different conditions. For example, it is often necessary to run large numbers of participants in order to ensure that ratings are statistically reliable. Furthermore, because each participant contributes a small number of ratings over the course of an entire experiment, researchers are often forced to conduct less powerful between-subject designs. Indeed, there would be little advantage to running multiple session, within-subject designs, as participants would need to spend an unreasonable number of hours in the laboratory in order for researchers to obtain reliable, within-subject estimates of performance.

The primary purpose of the present paper is to describe a novel contingency judgement procedure

constructed to avoid previous limitations due to task length. We reduced the length of a block of trials from several minutes to a few seconds using a streaming method involving the rapid sequential presentation of cue–outcome pairs, telescoping an entire block of trials into a single (streamed) trial. A presentation stream is depicted schematically in Figure 1.

The cue and the outcome are coloured geometric forms (a blue square as the cue and a red circle as the outcome) presented on a grey frame in the centre of a black monitor screen. Each 100-ms presentation consists of one of the four cue–outcome combinations (see inset), and presentations are separated by a 100-ms black screen.

The streaming procedure affords several advantages over previous procedures, and we elaborate on how the procedure may open interesting avenues for future research in the General Discussion. At the same time, it was important to establish the validity of the streaming-trial technique. To this end, we replicated two of the central findings in the contingency judgement literature using our streamed-trials technique. First, we establish that participant’s ratings of contingency correlated highly with the programmed (nominal) ΔP values used to describe the cue–outcome pairings for each presentation stream. Second, we investigated whether participant’s rating of contingency would be influenced by an

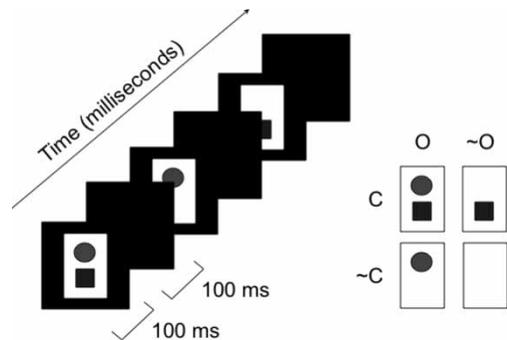


Figure 1. Left: a schematic illustrating the structure of a stream. Right: the four frames used in each. Squares are cues (C) and were presented in blue. Circles are outcomes (O) and were presented in red.

outcome density manipulation. In addition, we further probed the streamed-trials technique by including a frequency estimate judgement (see Wasserman & Shaklee, 1984). Specifically, after viewing each stream, either participants were asked for a contingency rating, or participants were asked to estimate the frequency of occurrence for each of the cue–outcome events (A, B, C, D) presented during the stream. Importantly, we were able to address these issues by manipulating two levels of contingency (noncontingent: $\Delta P = 0$, vs. contingent: $\Delta P = .467$), two levels of outcome density (high vs. low), and two levels of judgement type (contingency rating vs. frequency rating) using a completely within-participants design.

a

Participants

Participants were 37 students from undergraduate psychology classes at McMaster University who were given course credit for participation.

Apparatus and stimuli

The experiment was controlled by Apple Emac computers with 17" CRT displays running customized METACARD software. Participants sat at a viewing distance of approximately 60 cm from the computer screen. During each trial, participants viewed a stream of 60 cue–outcome pairs. Each cue–outcome pair was presented in a frame (6.4 cm in height \times 5 cm in width), displayed in the centre of the screen in grey against a black 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 1. When the cue was presented it was centred at the bottom of the frame. The cue was always a blue square (1.6 cm in height and width). When the outcome was presented it was centred at the top of the frame. The outcome was always a red circle (1.6 cm in diameter).

Design and procedure

We employed a $2 \times 2 \times 2$ within-subjects design involving contingency (noncontingent: $\Delta P = 0$, vs. contingent: $\Delta P = .467$), outcome density (high vs. low), and judgement type (contingency rating

vs. frequency estimate). The outcome probability for the noncontingent streams was .2 for the low outcome density condition and .8 for the high outcome density condition. The outcome probability for the contingent streams was .33 for the low outcome density condition and .67 for the high outcome density condition. The contingency and outcome density conditions were manipulated in four separate blocks, order counterbalanced across participants. Within each block we presented 20 trials, each consisting of a stream of 60 frames of cue–outcome pairs. Each stream was randomly generated from one of the four 2×2 contingency matrices (previously employed by Allan et al., 2005) displayed in Figure 2.

Each stream involved the rapid serial visual presentation of 60 frames depicting one of the four possible cue–outcome events. Each frame was displayed for 100 ms, and the stimulus-onset asynchrony (SOA) between frames was 100 ms. In total, each stream lasted approximately 12 s. The screen remained black during the inter-frame interval.

Following each stream participants made one of two judgements. Participants were not told in

	Low		High		
$\Delta P=0$	O	~O	$\Delta P=0$	O	~O
C	6	24	C	24	6
~C	6	24	~C	24	6
$\Delta P=.467$	O	~O	$\Delta P=.467$	O	~O
C	17	13	C	27	3
~C	3	27	~C	13	17

F 2. The top two matrices outline the frequency of the four cue–outcome pairings in a (left) low outcome density noncontingent stream, $\Delta P = 0$, $P(O) = .2$, and a (right) high outcome density noncontingent stream, $\Delta P = 0$, $P(O) = .8$. The bottom two matrices define a (left) low outcome density contingent stream, $\Delta P = .467$, $P(O) = .33$, and (right) high outcome density contingent stream, $\Delta P = .467$, $P(O) = .67$.

advance of watching each stream which judgement would be required. Instead, during each block of 20 streaming trials, participants randomly completed 10 contingency rating judgements and 10 frequency estimate judgements. Contingency ratings were assessed using a continuous scrollbar that participants could vary between a maximum negative value (-100) and a maximum positive value (+100). Frequency estimates were assessed by presenting participants with four pictures depicting each of the four possible cue-outcome events. An empty field was placed beside each picture, and participants were instructed to give an estimate of the frequency of occurrence for each cell presented in the previous stream (see Figure 3).

After each judgement participants were required to click a button on the computer monitor to initiate the next trial stream.

To summarize, every participant provided data for each of the four experimental conditions defined by two contingency values and two outcome density values, with order counterbalanced. For each condition, there were 20 streamed trials, each consisting of 60 presentations of cue-outcome pairs. On half the trials, a contingency rating was required, and on half the trials a frequency estimate was required, with order randomized.

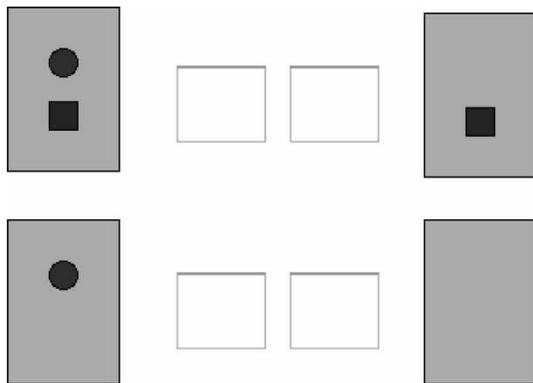


Figure 3. Depicts the display shown to participants for frequency estimate judgements. Each cue-outcome pair is shown beside an empty text field where participants could input their estimated frequency of occurrence for each cell. The displays presented to participants were coloured, with squares displayed in blue and circles displayed in red.

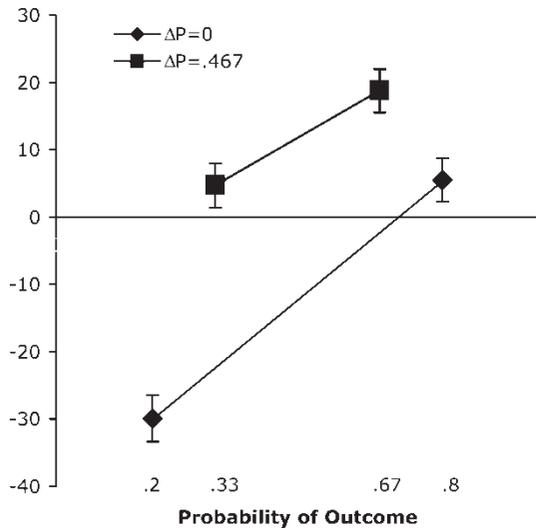
At the beginning of the experiment instructions appeared on the computer monitor. Participants were told that they would be watching streams of flashing squares and circles. Following each stream a test screen would appear, probing their knowledge of the stream in one of two ways. First, participants were instructed that the contingency rating should reflect their assessment of the strength of association between the square and the circle. It was further described that a positive rating should be given when the stream contained several trials in which the square was simultaneously presented with the circle (cue/outcome), as well as several trials in which no square or circle was presented (no cue/no outcome). Furthermore, it was explained that negative ratings should be given for sequences that contained several trials in which the square was presented without the circle (cue/no outcome) or vice versa (no cue/outcome). Finally, participants were instructed that a contingency rating of 0 would be appropriate if neither of the previous patterns was observed in the sequence. Second, in the case of frequency estimates, participants were shown an example screen displaying how frequency estimates for each cue-outcome pair were to be recorded and were further instructed that a frequency estimate should be given for each possible cue-outcome pair presented in a stream. Furthermore, during the instruction phase, example streams were shown to each participant. Specifically, each participant was shown one noncontingent stream and one contingent stream, labelled “zero contingency” and “positive contingency”, with example stream order counterbalanced across participants.



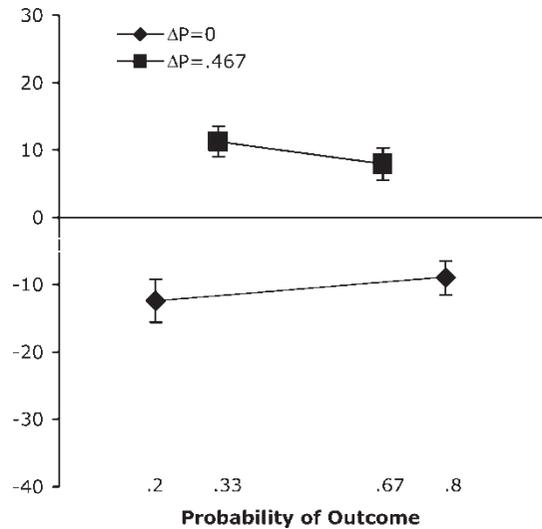
Contingency ratings

Each participant provided 10 contingency ratings under each of the four experimental conditions. The mean contingency rating, collapsed across participants, is displayed in Figure 4 for each of the four conditions.

For the noncontingent streams, mean ratings were -29.9 and 5.6 in the low and high



F 4. Mean contingency ratings (with standard error) as a function of ΔP (0 vs. .467) and outcome density.



F 5. Mean transformed ΔP scores derived from participant's frequency estimates (with standard error) as a function of ΔP (0 vs. .467) and outcome density.

outcome density conditions, respectively. For the contingent streams, mean ratings were 4.7 and 18.8 in the low and high outcome density conditions, respectively. These data were submitted to a 2×2 within-participants analysis of variance (ANOVA) with contingency (noncontingent vs. contingent), and outcome density (high vs. low) as factors. The main effect of contingency was significant, $F(1, 36) = 47.84$, $MSE = 443.59$, $p < .001$. Ratings were higher for contingent streams (11.8) than noncontingent streams (-12.2). The main effect of outcome density was significant, $F(1, 36) = 52.76$, $MSE = 429.89$, $p < .001$. Ratings were higher for high outcome density streams (12.2) than low outcome density streams (-12.6). There was also a significant Contingency \times Outcome Density interaction, $F(1, 36) = 15.73$, $p < .001$. The difference

in the ratings between the high and low outcome density conditions was larger for noncontingent streams (35.5) than contingent streams (14.05).¹

Frequency estimates

The frequency estimates from each participant were collapsed across each cell in the design and were transformed into ΔP values using the ΔP rule (Equation 1), and the mean transformed ΔP scores ($\times 100$) are displayed in Figure 5.

For the noncontingent streams, mean transformed ΔP scores were -12.4 and -9.0 in the low and high outcome density conditions, respectively. For the contingent streams, mean transformed ΔP score were 11.3 and 7.9 in the low and high outcome density conditions, respectively. These data were submitted to 2×2 within-participants ANOVA with contingency

¹ It is important to note that the interaction between outcome density and contingency cannot be interpreted due to an experimental confound between outcome density and contingency. The confound owed to the fact that the probability of an outcome was lower in the low outcome density, noncontingent condition (.2) than in the corresponding contingent condition (.33); similarly, the probability of an outcome was higher in the high outcome density, noncontingent condition (.8) than in the corresponding contingent condition (.67). The 2×2 contingency matrices containing this experimental confound were previously employed by Allan et al. (2005) and were employed in the current research to ensure that the streamed-trials procedure was as similar as possible to previous research.

(noncontingent vs. contingent) and outcome density (high vs. low) as factors. The main effect of contingency was significant, $F(1, 36) = 46.8$, $MSE = 326.07$, $p < .001$. The transformed ΔP scores were higher for contingent streams (9.61) than noncontingent streams (-10.7). The main effect of outcome density was not significant, $F(1, 36) < 1$, $MSE = 169.9$. Also, the Contingency \times Outcome Density interaction was not significant, $F(1, 36) = 3.18$, $MSE = 133.8$, $p = .08$.

Discussion

The purpose of the present experiment was to evaluate a novel streaming-trial method for measuring contingency judgements. The results are remarkably clear, replicating two central findings in the contingency judgement literature. First, participant's contingency ratings tracked the actual contingency between cue-outcome events presented during each stream. Second, participant's contingency ratings were influenced by outcome density. We argue that these findings adequately establish the streaming procedure as a viable method for measuring knowledge of contingency.

The frequency estimate manipulation also revealed a noteworthy finding. The transformed ΔP values derived from the frequency estimates demonstrated sensitivity to the contingency manipulation, but were not biased by the outcome density manipulation. This result is interesting for several reasons. First, we have established that the streaming-trials procedure can be probed with at least two dependent measures. Second, the frequency estimates provide converging evidence for Allan et al.'s (2005) claim that outcome density effects in ratings reflect a response bias and do not reflect participant's actual sensitivity to the contingency structure learned for each sequence. If the outcome density effect reflected changes in sensitivity due to differing levels of outcome density, then the straightforward expectation should be that other measures sensitive to contingency should also reveal an outcome density effect. Instead, it appears that

response biases may be introduced by the nature of the judgement question used to probe participants knowledge of the contingency structure.

Given this result, it is interesting to speculate on the possibility that outcome density effects belong to the larger class of statistical reasoning cognitive illusions. For example, research into statistical reasoning abilities (Kahneman & Tversky, 1996) has demonstrated that systematic biases in statistical estimates can arise from the format in which information is presented or tested. Consider the following Bayesian inference task that requires participants to estimate the likelihood that a woman has breast cancer given that her mammogram tested positive. Participants presented with the relevant base rate information in single-event probability format demonstrated a base-rate neglect effect and systematically overestimated the probability of having breast cancer. On the other hand, participants presented with the relevant base rate information in frequency format appeared to integrate base-rate knowledge into their judgements and were more accurate in their estimates of the number of women that would test positive (Gigerenzer & Hoffrage, 1995). Along these lines, an interesting avenue for future research would be to investigate the extent to which outcome density effects reflect a similar cognitive illusion mediated by the format in which contingency judgements are obtained.

Implications for future research

Research into human contingency judgements has generated an invaluable set of empirical findings and has spurred several major theoretical developments. Most notably, the contingency judgement literature has offered productive grounds for understanding the relationship between processes guiding animal and human learning. For example, a major line of inquiry has been to establish the extent to which human contingency judgements can be understood in terms of associative learning models (e.g., R-W model; Rescorla & Wagner, 1972) conventionally used to explain animal behaviour in conditioning contexts. In a more recent theoretical development, Allan et al. (2005), suggested that the processes underlying contingency judgements in

humans can be described in terms of signal detection theory (SDT). Allan et al. further argued that their signal detection analyses of contingency judgements should not necessarily be taken as an alternative to an associative interpretation; instead, they argued that SDT could serve as a more effective tool for evaluating association learning in the context of contingency judgement tasks.

One potential limitation of applying SDT to analysing contingency judgement tasks is that SDT is often employed in the context of psychophysical procedures that demand extensive within-subject measures of performance. Although conventional procedures are too lengthy for rigorous psychophysical testing, the streaming method developed here would lend itself handily toward direct psychophysical investigations. Indeed, one line of ongoing research has been to translate our streamed-trials procedure into a psychophysical design whereby trained participants are asked to discriminate between streams of varying contingency across multiple training sessions. Using this approach it is possible to generate psychometric functions describing each participant's sensitivity to contingent information across varying levels of contingency. In other words, it is possible to estimate each participant's just-noticeable-difference between two levels of contingency. This psychophysical approach can be used to directly investigate Allan et al.'s (2005) claim that outcome density effects are driven by response biases. For example, the response-bias account would predict that the slope of a psychometric function describing a participant's sensitivity to contingency will not vary with outcome density; instead the entire psychometric function should shift as function of the participant's response bias.

More generally, the SDT approach to analysing contingency judgements provides a unique tool for determining factors that separately influence participant's sensitivity as well as response bias. Estimates of these separate measures may further clarify the processes mediating performance across several areas of interest in the contingency judgement literature. For example, apart from outcome density manipulations, contingency

judgements can be influenced by a range of processes from those susceptible to classical conditioning phenomena (e.g., blocking and several others; see Allan, 1993) to processes responsible for framing information in terms of causal models (De Houwer & Beckers, 2002). It would be useful to interpret these influences in terms of SDT, as the measures of sensitivity and bias would provide unique insight into the nature of the processes mediating performance. For example, would blocking manipulations uniquely influence measures of sensitivity? The influence of a blocking manipulation could be measured using the streamed-trials procedure by introducing another cue (e.g., a triangle), of varying cue validity, to the stream of trials. Would the influence of causal models uniquely influence measures of response bias? It would be interesting to determine whether participant's ratings of contingency would be influenced by varying a cover story aimed at manipulating participants understanding of the causal relationship between cues and outcomes presented during a stream of trials. More generally, what factors increase sensitivity and sway response bias in the context of contingency judgements? We are optimistic that some insight into these questions can be obtained by combining SDT with a psychophysical approach to measuring contingency judgements using the streamed-trials procedure.

In sum, the purpose of this paper was to establish a novel, streamed-trials procedure for measuring contingency judgements. Using the streamed-trials procedure, we replicated two central findings in the contingency judgement literature. Specifically, participant's ratings of contingency tracked the objective ΔP values, as well as the outcome density manipulation for each stream of cue-outcome events. Furthermore, we demonstrated that the outcome density effect depended on the information format of the question used to probe participant's knowledge of the contingency structure of each stream. This result lends further support to the notion that outcome density effects are mediated by biases initiated at the time of response. Taken together, we have established that the streamed-trials method is a reliable

procedure for measuring knowledge of contingency, and we are optimistic that this procedure can be further augmented to provide many sources of insight into the processes responsible for detecting and interpreting contingencies between events.

Original manuscript received 26 May 2006

Accepted revision received 30 October 2006

First published online 4 April 2007



- Allan, L. G. (1980). A note on measurement of contingency between two binary variables in judgment tasks. *Bulletin of the Psychonomic Society*, 15, 147–149.
- Allan, L. G. (1993). Human contingency judgments: Rule based or associative? *Psychological Bulletin*, 114, 435–448.
- Allan, L. G., & Jenkins, H. M. (1983). The effect of representations of binary variables on judgment of influence. *Learning and Motivation*, 14, 381–405.
- Allan, L. G., Siegel, S., & Tangen, J. M. (2005). A signal detection analysis of contingency data. *Learning and Behaviour*, 33, 250–263.
- Alloy, L. B., & Abramson, L. Y. (1979). Judgments of contingency in depressed and nondepressed students. Sadder but wiser? *Journal of Experimental Psychology: General*, 108, 441–485.
- De Houwer, J., & Beckers, T. (2002). A review of recent developments in research and theory on human contingency learning research. *Quarterly Journal of Experimental Psychology*, 55B, 289–310.
- Dickinson, A., Shanks, D., & Evenden, J. (1984). Judgment of act–outcome contingency: The role of selective attribution. *Quarterly Journal of Experimental Psychology*, 36A, 29–50.
- Gallistel, C. R. (2002). Frequency, contingency and the information processing theory of conditioning. In P. Sedlmeier & T. Betsch (Eds.), *Frequency processing and cognition*. Oxford, UK: University Press.
- Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review*, 102, 684–704.
- Gluck, M. A., & Bower, G. H. (1988). From conditioning to category learning. *Journal of Experimental Psychology: General*, 117, 227–247.
- Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. New York: Wiley.
- Kahneman, D., & Tversky, A. (1996). On the reality of cognitive illusions. *Psychological Review*, 103, 582–591.
- Kao, S. E., & Wasserman, E. A. (1993). Assessment of an information integration account of contingency judgment with examination of subjective cell importance and method of information presentation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 19, 1363–1386.
- Mandel, D., & Lehman, D. (1998). Integration of contingency information in judgments of cause, covariation, and probability. *Journal of Experimental Psychology: General*, 127, 269–285.
- Perales, J., Catena, A., Gonzalez, J., & Shanks, D. (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.
- Rescorla, R. A., & Wagner, A. R. (1972). A theory of Pavlovian conditioning: Variations in the effectiveness of reinforcement and non-reinforcement. In A. H. Black & W. F. Prokossy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–99). New York: Appleton-Century-Crofts.
- Shanks, D. R. (1985). Continuous monitoring of human contingency judgment. *Quarterly Journal of Experimental Psychology*, 37B, 1–21.
- Shanks, D. R. (1986). Selective attribution and the judgment of causality. *Learning and Motivation*, 17, 311–334.
- Wasserman, E. A., Chatlosh, D. L., & Neunaber, D. J. (1983). Perceptions of causal relations in humans: Factors affecting judgments of response–outcome contingencies under free-operant procedures. *Learning and Motivation*, 14, 406–432.
- Wasserman, E. A., & Shaklee, H. (1984). Judging response–outcome relations: The role of response–outcome contingency, outcome probability, and method of information presentation. *Memory & Cognition*, 12, 170–286.

Copyright of Quarterly Journal of Experimental Psychology is the property of Psychology Press (UK) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.